

GenCore version 5.1.3

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:03:23 ; Search time 17 Seconds
(without alignments)

378.883 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGGGGLGGKGCPSNEIFSR.....CRIGYLRNKKKVCVPRSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: piri:*

2: piri:*

3: piri:*

4: piri:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	Length	DB	ID	Description
1	133	34.0	245	2	T21654		hypothetical prote
2	118	30.2	561	2	T27318		hypothetical prote
3	118	30.2	626	2	T27319		hypothetical prote
4	113	28.9	2155	2	T30197		alpha tectorin - m
5	113	28.9	5376	2	T42215		zonadhesin - mouse
6	111.5	28.5	166	2	H89044		protein B0238.12 [
7	109.5	28.0	802	2	T24293		hypothetical prote
8	109.5	28.0	949	2	T24294		hypothetical prote
9	107.5	27.5	2476	2	T34022		zonadhesin - pig
10	100.5	25.7	137	2	T15609		hypothetical prote
11	100	25.6	140	2	T16574		hypothetical prote
12	98	25.1	2813	1	VWU		hypothetical prote
13	97.5	24.9	2120	2	T30243		von Willebrand fac
14	96.5	24.7	1036	2	T17405		alpha tectorin - c
15	94.5	24.2	98	2	C89046		scavenger receptor
16	94.5	24.2	1023	2	T30257		protein C1068.4 [i
17	93	23.8	192	2	T25513		IgC Fc binding pro
18	92.5	23.7	195	2	T28803		hypothetical prote
19	92	23.5	249	2	T24604		hypothetical prote
20	91	23.3	1373	2	J60095		hypothetical prote
21	90.5	23.1	135	2	T15610		gastric mucin MUC5
22	90.5	23.1	3133	2	S52093		hypothetical prote
23	89.5	22.9	1642	2	T19130		hemocytin - silkwo
24	88.5	22.6	145	2	T15608		hypothetical prote
25	86	22.0	1700	2	S08167		hypothetical prote
26	84.5	21.6	780	2	A34102		Balbani ring 3 pr
27	84.5	21.6	1513	2	A54895		von Willebrand fac
28	83.5	21.4	13288	2	T03099		mucin 2, intestinal
29	80.5	20.6	63	2	S07127		chymotrypsin/elast

ALIGNMENTS

RESULT 1

T21654

hypothetical protein F32D8.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T21654

R:Wilkinson, J.

submitted to the EMBL Data Library, June 1996

A:Reference number: Z19454

A:Accession: T21654

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-245 <WIL>

A:Cross-references: EMBL:Z74031; PIDN:CAA98455.1; GSPDB:GN00023; CESP:F32D8.3

A:Experimental source: clone F32D8

C:Genetics:

A:Gene: CESP:F32D8.3

A:Map position: 5

A:Introns: 61/1; 83/3

C:Superfamily: Caenorhabditis elegans hypothetical protein F32D8.3

Query Match 34.0%; Score 133; DB 2; Length 245;

Best Local Similarity 43.1%; Pred. No. 4.4e-06;

Matches 28; Conservative 7; Mismatches 20; Indels 10; Gaps 4;

QY 6 LGGRGKCPSEIFSRCDGRQRCRCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVCV---- 61

Db 57 INGEENCNPMFQSFSSHC--ACESTCNN--PDYCSK-CEPGCTCRNGFVRNSIKLCVLP EE 111

QY 62 -PRSK 65

Db 112 CPTK 116

RESULT 2

T27318

hypothetical protein Y69H2.3a - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27318

R:McMurray, A.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z20343

A:Accession: T27318

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-561 <WIL>

A:Cross-references: EMBL:Z98877; PIDN:CA54472.1; GSPDB:GN00023; CESP:Y69H2.3a

A:Experimental source: clone Y69H2

C:Genetics:

A:Gene: CESP:Y69H2.3a

A:Map position: 5

30 80.5 20.6 99 2 S44658 ZK353.2 protein -

31 80 20.5 1371 2 A33837 insulin-like growt

32 79.5 20.3 100 2 T23389 hypothetical prote

33 79.5 20.3 1506 2 T30886 integumentary muc

34 78 19.9 490 2 T32003 hypothetical prote

35 77.5 19.8 63 2 S08572 chymotrypsin/elast

36 77.5 19.8 915 2 T21772 hypothetical prote

37 77.5 19.8 927 2 T21772 hypothetical prote

38 77 19.7 169 2 T15611 metallothionein 10

39 76.5 19.6 72 2 S39416 hypothetical prote

40 76.5 19.6 2195 2 T34264 hypothetical prote

41 75.5 19.3 253 2 T25768 hypothetical prote

42 75.5 19.3 348 2 T28623 hypothetical prote

43 75.5 19.3 349 2 D72175 G2R protein - vari

44 75.5 19.3 349 2 D36858 gene G2R protein -

45 75.5 19.3 2910 2 T42214 otogelin - mouse

A:Reference number: Z19870

A:Accession: T24293

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-802 <WIL>

A:Cross-references: EMBL:Z81110; PIDN: CAB03259.1; GSPDB: GN00023; CESP: T01D3.3a

A:Experimental source: clone T01D3

C:Genetics:

A:Gene: CESP:T01D3.3a

A:Map position: 5

A:Introns: 74/1; 121/1; 200/2; 493/1; 673/3; 772/2

Query Match

Best Local Similarity 28.0%; Score 109.5; DB 2; Length 802;

Matches 23; Conservative 7; Mismatches 24; Indels 5; Gaps 3;

QY 12

CPSNEIFSRCDRCQRCFCPNVVPKPLC-ITKAPGCVRLGYLR---NKKKVCVPRSKC 66

Db 82 CGANEQYSACFSQCPSCQD-PSTPACPAPGCGICLPGYIRRDSSPRACVPRGLC 139

RESULT 8

T24294

hypothetical protein T01D3.3b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T24294

R:Steward, C.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19870

A:Accession: T24294

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-949 <WIL>

A:Cross-references: EMBL:Z81110; PIDN: CAB03260.1; GSPDB: GN00023; CESP: T01D3.3b

A:Experimental source: clone T01D3

C:Genetics:

A:Gene: CESP:T01D3.3b

A:Map position: 5

A:Introns: 74/1; 126/1; 161/2; 221/1; 268/1; 347/2; 640/1; 820/3; 919/2

Query Match

Best Local Similarity 28.0%; Score 109.5; DB 2; Length 949;

Matches 23; Conservative 7; Mismatches 24; Indels 5; Gaps 3;

QY 12

CPSNEIFSRCDRCQRCFCPNVVPKPLC-ITKAPGCVRLGYLR---NKKKVCVPRSKC 66

Db 229 CGANEQYSACFSQCPSCQD-PSTPACPAPGCGICLPGYIRRDSSPRACVPRGLC 286

RESULT 9

T34022

zonadhesin - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000

C:Accession: T34022

J:Hardy, D.M.; Garbers, D.L.

J: Biol. Chem. 270, 26025-26028, 1995

A:Title: A sperm membrane protein that binds in a species-specific manner to the egg ext

A:Reference number: Z21464; MUID: 96064658; PMID: 7592795

A:Accession: T34022

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2476 <HAR>

A:Cross-references: EMBL:U40024; NID: g1066465; PID: g1066466; PIDN: AAC48486.1

A:Experimental source: strain Meishan; testis

C:Genetics:

A:Gene: zan

C:Function:

A:Description: may be involved in sperm adhesion to the zona pellucida

Query Match

Best Local Similarity 27.5%; Score 107.5; DB 2; Length 2476;

Matches 23; Conservative 7; Mismatches 24; Indels 5; Gaps 3;

Matches 22; Conservative 9; Mismatches 18; Indels 17; Gaps 3;

QY 11 KCPSEIFSRCDGRC-----QRCFNPVVPKPLCIKICAPGCVRLGYLRNKKKVCV 61

Db 1455 KCPGSSYSTCANPCPCTCLSLNNPSYCPSTLP-----CABGCECQKHILLSGTS-CV 1506

QY 62 PRSKCG 67

Db 1507 PLSOCG 1512

RESULT 10

T15609

hypothetical protein C25E10.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15609

R:Bradshaw, H.

submitted to the EMBL Data Library, February 1996

A:Description: The sequence of C. elegans cosmid C25E10.

A:Reference number: Z18376

A:Accession: T15609

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-137 <BRA>

A:Cross-references: EMBL:U50311; NID: g1226295; PID: g1226303; PIDN: AAA92313.1; CESP: C25E10.8

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:C25E10.8

A:Introns: 1/3; 19/1; 42/3

Query Match

Best Local Similarity 25.7%; Score 100.5; DB 2; Length 137;

Matches 21; Conservative 7; Mismatches 26; Indels 3; Gaps 3;

QY 11

KCPSEIFSRCDGRCQRCFNPVVPKPLCIKICAPG-CVCRILGYLRNKKKVCVPRSKC 66

Db 81 KCPENETFFRCGTACEPTCEKPGPRP-CTRQCIVNVQCSCSGFVRNGYR-CTELKEC 135

RESULT 11

T16574

hypothetical protein K05F1.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16574

R:Wohlmann, P.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid K05F1.

A:Reference number: Z18537

A:Accession: T16574

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-140 <WOH>

A:Cross-references: EMBL:U29377; NID: g868173; PID: g868180; PIDN: AAA68717.1; CESP: K05F1.10

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:K05F1.10

A:Introns: 31/3; 77/2

Query Match

Best Local Similarity 25.6%; Score 100; DB 2; Length 140;

Matches 21; Conservative 10; Mismatches 25; Indels 8; Gaps 2;

QY 11

KCPSEIFSRSC--DGRQRCFNPVVPKPLCI-----KICAPGCVRLGYLRNKKKVCV 62

Db 52 ECQKHHLICGPERHCDRTCELFSPPHCLNHLHAKCYFPRVCNDGVYRSEKIGICR 111

QY 63

RSKC 66

Db 112

PSHC 115

RESULT 12

VHUU

von Willebrand factor precursor - human

C:Species: Homo sapiens (man)

C:Date: 04-Dec-1986 #sequence-revision 30-Jun-1993 #text-change 22-Jun-1999

C:Accession: A34480; S02377; A37139; S23676; A25298; A25469; A25366; S23618; S23645; A94

R:Mancuso, D.J.; Tuley, E.A.; Westfield, L.A.; Worrall, N.K.; Shelton-Inloes, B.B.; Sorace, J.E.;

J. Biol. Chem. 264, 19514-19527, 1989

A:Title: Structure of the gene for human von Willebrand factor.

A:Reference number: A34480; MUID:90062044; PMID:2584182

A:Accession: A34480

A:Molecule type: DNA

A:Residues: 1-2813 <MAN>

A:Cross-references: EMBL:M25864

R:Bonthron, D.; Orkin, S.H.

Eur. J. Biochem. 171, 51-57, 1988

A:Title: The human von Willebrand factor gene. Structure of the 5' region.

A:Reference number: S02377; MUID:88111704; PMID:2828057

A:Accession: S02377

A:Molecule type: DNA

A:Residues: 1-177 <BO2>

A:Cross-references: EMBL:X06828

R:Mancuso, D.J.; Tuley, E.A.; Westfield, L.A.; Lester-Mancuso, T.L.; Le Beau, M.M.; Sorace, J.E.;

Biochemistry 30, 253-269, 1991

A:Title: Human von Willebrand factor gene and pseudogene: structural analysis and differentiation.

A:Reference number: A37139; MUID:91105089; PMID:1988024

A:Accession: A37139

A:Molecule type: DNA

A:Residues: 990-1947 <MAD>

A:Cross-references: GB:M60675; NID:g340357; PIDN:AAA61295.1; PID:g553810

A:Note: the authors translated the codon CGC for residue 156 as Gln

R:Collins, C.J.; Underdahl, J.P.; Levene, R.B.; Ravera, C.P.; Morin, M.J.; Dombalagian, R.;

Proc. Natl. Acad. Sci. U.S.A. 84, 4393-4397, 1987

A:Title: Molecular cloning of the human gene for von Willebrand factor and identification of the gene.

A:Reference number: S23676; MUID:87260814; PMID:3496594

A:Accession: S23676

A:Molecule type: DNA

A:Residues: 2731-2813 <COL>

A:Cross-references: EMBL:M16945

R:Bonthron, D.; Orr, E.C.; Mitsock, L.M.; Ginsburg, D.; Handin, R.I.; Orkin, S.H.

Nucleic Acids Res. 14, 7125-7127, 1986

A:Title: Nucleotide sequence of pre-pro-von Willebrand factor cDNA.

A:Reference number: A25298; MUID:87016349; PMID:3489923

A:Accession: A25298

A:Molecule type: mRNA

A:Residues: 1-470, 'V', 472-2813 <BON>

A:Cross-references: EMBL:X04385

R:Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.

EMBO J. 5, 1839-1847, 1986

A:Title: Full-length von Willebrand factor (vWF) cDNA encodes a highly repetitive protein.

A:Reference number: A91044; MUID:87004550; PMID:3019665

A:Accession: A25469

A:Molecule type: mRNA

A:Residues: 1-470, 'V', 472-483, 'R', 485-1022, 'K', 1024-1025, 'E', 1027-1400 <VER>

A:Cross-references: EMBL:X04146

A:Note: this sequence has been revised in reference A91056

R:Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H.

EMBO J. 5, 3074, 1986

A:Reference number: A91056

A:Accession: A25366

A:Molecule type: mRNA

A:Residues: 1021-1030 <VE2>

A:Note: this is a revision to the sequence from reference A91044

R:Shelton-Inloes, B.B.; Broze Jr., G.J.; Miletich, J.P.; Sadler, J.E.

Biochem. Biophys. Res. Commun. 144, 657-665, 1987

A:Title: Evolution of human von Willebrand factor: cDNA sequence polymorphisms, repeated

A:Reference number: S23618; MUID:87213253; PMID:3495266

A:Accession: S23618

A:Molecule type: mRNA

A:Residues: 1-120 <SH2>

A:Cross-references: EMBL:M17588; NID:g799330; PIDN:AAA65940.1; PID:g340316

A:Accession: S23645

A:Molecule type: protein

A:Residues: 23-56 <SH3>

R:Sadler, J.E.; Shelton-Inloes, B.B.; Sorace, J.M.; Harlan, J.M.; Titani, K.; Davie, J.

Proc. Natl. Acad. Sci. U.S.A. 82, 6394-6398, 1985

A:Title: Cloning and characterization of two cDNAs coding for human von Willebrand factor

A:Reference number: A94060; MUID:86016708; PMID:2864688

A:Accession: A94060

A:Molecule type: mRNA

A:Residues: 'WA', 739, 'C', 744-769, 'H', 771-788, 'A', 790-803, 'S', 805-873; 1289-1471, 'D', 14

R:Shelton-Inloes, B.B.; Titani, K.; Sadler, J.E.

Biochemistry 25, 3164-3171, 1986

A:Title: cDNA sequences for human von Willebrand factor reveal five types of repeated

A:Reference number: A90504; MUID:86269894; PMID:3488076

A:Accession: A90504

A:Molecule type: mRNA

A:Residues: 781-788, 'A', 790-1424 <SHE>

A:Note: 852-Gln, 857-Asp, and 1381-Thr were also found

R:Ginsburg, D.; Handin, R.I.; Bonthron, D.T.; Donlon, T.A.; Bruns, G.A.P.; Latt, S.A.

Science 228, 1401-1406, 1985

A:Title: Human von Willebrand factor (vWF): isolation of complementary DNA (cDNA) clob

A:Reference number: A44178; MUID:85244588; PMID:3874428

A:Accession: A44178

A:Molecule type: mRNA

A:Residues: 2621-2813 <GIN>

A:Cross-references: EMBL:X03028; NID:g340308; PIDN:AAA61293.1; PID:g340309

R:Verweij, C.L.; de Vries, C.J.M.; Distel, B.; van Zonneveld, A.J.; van Kessel, A.G.;

Nucleic Acids Res. 13, 4699-4717, 1985

A:Title: Construction of cDNA coding for human von Willebrand factor using antibody p

A:Reference number: S07363; MUID:85369603; PMID:3875078

A:Accession: S07363

A:Molecule type: mRNA

A:Residues: 2731-2813 <VE3>

A:Cross-references: EMBL:X02672; NID:g37939; PIDN:CAA26503.1; PID:g37940

R:Lynch, D.C.; Zimmerman, T.S.; Collins, C.J.; Brown, M.; Morfin, M.J.; Ling, E.H.; Li

Cell 41, 49-56, 1985

A:Title: Molecular cloning of cDNA for human von Willebrand factor: authentication by

A:Reference number: S23678; MUID:85201687; PMID:3873280

A:Accession: S23678

A:Molecule type: mRNA

A:Residues: 2731-2813 <LYN>

A:Cross-references: EMBL:X03028

R:Titani, K.; Kumar, S.; Takio, K.; Ericsson, L.H.; Wade, R.D.; Ashida, K.; Walsh, K.

Biochemistry 25, 3171-3184, 1986

A:Title: Amino acid sequences of human von Willebrand factor.

A:Reference number: A90505; MUID:86269895; PMID:3524673

A:Accession: A90505

A:Molecule type: protein

A:Residues: 764-788, 'A', 790-1471, 'D', 1473-2813 <TIT>

A:Note: 789-Thr was also found

R:Chopek, M.W.; Girma, J.P.; Fujikawa, K.; Davie, E.W.; Titani, K.

Biochemistry 25, 3146-3155, 1986

A:Title: Human von Willebrand factor: a multivalent protein composed of identical sub

A:Reference number: A23464; MUID:86269892; PMID:3015199

A:Accession: A23464

A:Molecule type: protein

A:Residues: 764-773; 2803-2813 <CHO>

R:Dent, J.A.; Berkowitz, S.D.; Ware, J.; Kasper, C.K.; Ruggeri, Z.M.

Proc. Natl. Acad. Sci. U.S.A. 87, 6306-6310, 1990

A:Title: Identification of a cleavage site directing the immunohemical detection of

A:Reference number: A36013; MUID:90349604; PMID:2385594

A:Accession: A36013

A:Molecule type: protein

A:Residues: 1606-1617 <DEN>

R:Ray, P.J.; Kawai, Y.; Wagner, D.D.; Ginsburg, D.; Bonthron, D.; Ohlsson-Wilhelm, B.

Science 232, 995-998, 1986

A:Title: Propolypeptide of von Willebrand factor circulates in blood and is identical

A:Reference number: A60913; MUID:86208144; PMID:3486471

A:Accession: A60913

A:Molecule type: protein

A:Residues: 576-590 <FAY>

C:Genetics:

A:Gene: GDB:VWF

A:Cross-references: GDB:119125; OMIM:193400

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:03:22 ; Search time 29 Seconds
(without alignments)
95.825 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GFGGLGGRGKCPSEIFSR.....CRILGLRNKKKVCVPRSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	113	28.9	5376	1 ZAN_MOUSE	O88799 sus musculus
2	109	27.9	2700	1 ZAN_HUMAN	Q9y493 homo sapien
3	107.5	27.5	2476	1 ZAN_PIG	Q28983 sus scrofa
4	104.5	26.7	56	1 AMCI_APIME	P56682 apis mellif
5	103.5	26.5	2813	1 VWF_CANFA	Q28295 canis famil
6	101.5	26.0	937	1 VWF_BOVIN	P80012 bos taurus
7	98.5	25.2	2282	1 ZAN_RABIT	P57999 oryctolagus
8	98	25.1	115	1 A62F_DROME	O46202 drosophila
9	98	25.1	2813	1 VWF_HUMAN	P04275 homo sapien
10	91.5	23.4	2482	1 VWF_PIG	Q28833 sus scrofa
11	90.5	23.1	3133	1 HMCT_BOMMO	P98092 bombyx mori
12	86	22.0	1700	1 BAR3_CHITE	Q03376 chironomus
13	80.5	20.6	63	1 ICE1_ASCSU	P07851 ascaris suu
14	80.5	20.6	99	1 YOJ2_CAEEL	P34625 caenorhabdi
15	80	20.5	1370	1 IGR1_RAT	P24062 rattus norv
16	80	20.5	1373	1 IGR1_MOUSE	Q60751 mus musculus
17	79.5	20.3	351	1 CRMB_COMPX	O73559 corpopo viru
18	78.5	20.1	65	1 ICE2_ASCSU	P07852 ascaris suu
19	76.5	19.6	72	1 MT14_MYTED	P80246 mytilus edu
20	76.5	19.6	72	1 MT1B_MYTED	O62554 mytilus edu
21	76.5	19.6	349	1 CRMB_CAMPS	Q8uyav7 camelpox vi
22	76.5	19.6	855	1 ST14_MOUSE	P56677 mus musculus
23	75.5	19.3	349	1 CRMB_VARV	P34015 variola vir
24	75.5	19.3	5179	1 MUC2_HUMAN	Q02817 homo sapien
25	75	19.2	956	1 TSP3_MOUSE	Q05895 mus musculus
26	75	19.2	4289	1 TENX_HUMAN	P22105 homo sapien
27	73	18.7	60	1 MT_PERPL	P52725 perca fluvi
28	72.5	18.5	1416	1 YN81_CAEEL	Q03610 caenorhabdi
29	72	18.4	326	1 VT2_MXAVL	P29825 myxoma viru
30	72	18.4	956	1 TSP3_HUMAN	P49746 homo sapien
31	72	18.4	1367	1 IGR1_HUMAN	P08069 homo sapien
32	71	18.2	1746	1 TENA_PIG	Q29116 sus scrofa
33	70.5	18.0	62	1 ITR1_ASCSU	P19398 ascaris suu

34 70.5 18.0 72 1 MT12_MYTED
35 69 17.6 60 1 MTB_ONCMY
36 69 17.6 60 1 MT_ESOLU
37 69 17.6 77 1 PIF3_BOMMO
38 69 17.6 1339 1 ERB3_RAT
39 68.5 17.5 72 1 MT14_MYTED
40 68 17.4 60 1 MT_PLEPL
41 68 17.4 60 1 MT_PSEAM
42 68 17.4 4660 1 LRF2_RAT
43 67.5 17.3 182 1 KRUC_SHEEP
44 67 17.1 62 1 MT4_HUMAN
45 67 17.1 328 1 YRP1_CAEEL

P80247 mytilus edu
P09862 oncorhynchu
P25127 esox lucius
Q10731 bombyx mori
Q62799 rattus norv
P80249 mytilus edu
P07216 pleuronecte
P55945 pseudopleur
P98158 rattus norv
P26372 ovis aries
P47944 homo sapien
Q10043 caenorhabdi

ALIGNMENTS

RESULT 1
ZAN_MOUSE STANDARD; PRT; 5376 AA.
AC O88799; O08647;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zonadhesin precursor.
GN ZAN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=98123114; PubMed=9452463;
RA Gao Z., Garbers D.L.;
RT "Species diversity in the structure of zonadhesin, a sperm-specific
RT membrane protein containing multiple cell adhesion molecule-like
RT domains".
RL J. Biol. Chem. 273:3415-3421(1998).
RN [2]
RP SEQUENCE OF 4864-5376 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=97271566; PubMed=9126492;
RA Gao Z., Harumi T., Garbers D.L.;
RT "Chromosome localization of the mouse zonadhesin gene and the human
RT zonadhesin gene (ZAN)".
RL Genomics 41:119-122(1997).
CC -!- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
CC SIGNALING.
CC -!- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
CC APICAL REGION OF THE SPERM HEAD.
CC -!- TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS.
CC -!- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATES SPERM ADHESION TO THE
CC ZONA PELLUCIDA.
CC -!- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,
CC THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF
CC SPERMATOZOA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.
CC -!- DOMAIN: THE VWFD DOMAIN 2 MAY MEDIATE COVALENT
CC OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).
CC -!- SIMILARITY: CONTAINS 3 MAM DOMAINS.
CC -!- SIMILARITY: CONTAINS 25 VWFD DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U97068; AAC26680.1; -.
 DR EMBL; U83190; AAC53125.1; -.
 DR MGD; MGI:106656; Zan.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR003645; FOLN.
 DR InterPro; IPR000998; MAM_domain.
 DR InterPro; IPR002919; TIL_Cysrich.
 DR InterPro; IPR003328; TILa_Cysrich.
 DR InterPro; IPR001007; VWF_C.
 DR InterPro; IPR001846; VWF_D.
 DR Pfam; PF00094; vwd; 4.
 DR Pfam; PF00629; MAM; 3.
 DR Pfam; PF01826; TIL; 25.
 DR Pfam; PF02345; TILa; 25.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00274; FOLN; 11.
 DR SMART; SM00137; MAM; 2.
 DR SMART; SM00214; VWC; 4.
 DR SMART; SM00216; VWD; 4.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 18.
 DR PROSITE; PS00740; MAM_1; FALSE_NEG.
 DR PROSITE; PS00060; MAM_2; 3.
 KW Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
 Repeat.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 5376 ZONADHESIN.
 FT DOMAIN 18 5310 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 5311 5337 POTENTIAL.
 FT DOMAIN 5338 5376 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 45 210 MAM 1.
 FT DOMAIN 215 374 MAM 2.
 FT DOMAIN 377 542 MAM 3.
 FT DOMAIN 547 1170 80 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
 (MUCIN-LIKE DOMAIN).
 FT DOMAIN 1171 1280 WVED 1 (PARTIAL).
 FT DOMAIN 1281 1669 WVED 2.
 FT DOMAIN 1670 2056 WVED 3.
 FT DOMAIN 2057 2459 WVED 4.
 FT DOMAIN 2460 2579 WVED 5 (PARTIAL).
 FT DOMAIN 2580 2699 WVED 6 (PARTIAL).
 FT DOMAIN 2700 2819 WVED 7 (PARTIAL).
 FT DOMAIN 2820 2939 WVED 8 (PARTIAL).
 FT DOMAIN 2940 3059 WVED 9 (PARTIAL).
 FT DOMAIN 3060 3179 WVED 10 (PARTIAL).
 FT DOMAIN 3180 3299 WVED 11 (PARTIAL).
 FT DOMAIN 3300 3416 WVED 12 (PARTIAL).
 FT DOMAIN 3417 3536 WVED 13 (PARTIAL).
 FT DOMAIN 3537 3656 WVED 14 (PARTIAL).
 FT DOMAIN 3657 3776 WVED 15 (PARTIAL).
 FT DOMAIN 3777 3892 WVED 16 (PARTIAL).
 FT DOMAIN 3893 4029 WVED 17 (PARTIAL).
 FT DOMAIN 4029 4148 WVED 18 (PARTIAL).
 FT DOMAIN 4149 4263 WVED 19 (PARTIAL).
 FT DOMAIN 4264 4383 WVED 20 (PARTIAL).
 FT DOMAIN 4384 4503 WVED 21 (PARTIAL).
 FT DOMAIN 4504 4623 WVED 22 (PARTIAL).
 FT DOMAIN 4624 4743 WVED 23 (PARTIAL).
 FT DOMAIN 4744 4863 WVED 24 (PARTIAL).
 FT DOMAIN 4864 5261 WVED 25.
 FT DOMAIN 5259 5295 EGF-LIKE.
 FT DISULFID 5263 5274 BY SIMILARITY.
 FT DISULFID 5268 5283 BY SIMILARITY.
 FT DISULFID 5285 5294 BY SIMILARITY.
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1216 1216 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1239 1239 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1314 1314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1814 1814 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1908 1908 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1933 1933 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2028 2028 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 2111 2111 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2142 2142 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2332 2332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2533 2533 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2575 2575 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2692 2692 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2812 2812 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3052 3052 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3065 3065 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3144 3144 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3172 3172 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3288 3288 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3292 3292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3782 3782 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4005 4005 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4136 4136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4243 4243 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4254 4254 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4335 4335 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4376 4376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 4586 4586 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 5136 5136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 5252 5252 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 5376 AA; 579908 MM; 0E44DB77DF2A2620 CRC64;
 Query Match 28.9%; Score 113; DB 1; Length 5376;
 Best Local Similarity 37.5%; Pred. No. 0.0034;
 Matches 24; Conservative 8; Mismatches 22; Indels 10; Gaps 3;
 QY 11 KCPSEIFSRCDGRCQFCFNVVPRPLC-----IKICAPGCVGLGVLRNKKVCVPR 63
 Db 4743 KCPANSLYTHCLPTCLPSCSN--PDGRCEGTSKAPSTCREGCVQPGYLLN-KDTCVHK 4799
 QY 64 SKCG 67
 Db 4800 NQCG 4803
 RESULT 2
 ID ZAN_HUMAN STANDARD; PRT; 2700 AA.
 AC Q9Y493; O00218;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Zonadhesin (Fragment).
 GN ZAN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-2379 FROM N.A.
 RX MEDLINE=99018118; PubMed=979793;
 RA Glockner G., Scherer S., Schattevoy R., Boright A., Weber J.,
 RA Tsui L.C., Rosenthal A.;
 RT "Large-scale sequencing of two regions in human chromosome 7q22:
 RT analysis of 650 kb of genomic sequence around the EPO and CUTLI loci
 RT reveals 17 genes.";
 RL Genome Res. 8:1060-1073(1998).
 RN [2]
 RP SEQUENCE OF 2338-2700 FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=97271566; PubMed=9126492;
 RA Gao Z., Harumi T., Garbers D.L.;
 RT "Chromosome localization of the mouse zonadhesin gene and the human
 RL zonadhesin gene (ZAN).";
 RL Genomics 41:119-122(1997).
 CC -!- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
 CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
 CC SIGNALING.
 CC -!- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE

DR SMART; SM00137; MAM; 1.
 DR SMART; SM00214; VWC; 2.
 DR SMART; SM00216; VWD; 4.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 4.
 DR PROSITE; PS00740; MAM_1; 1.
 DR PROSITE; PS00060; MAM_2; 2.
 KW Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion;
 Repeat.
 FT CHAIN 1 29
 FT CHAIN 30 2476
 FT DOMAIN 30 2418
 FT TRANSMEM 2419 2439
 FT DOMAIN 2440 2476
 FT DOMAIN 31 144
 FT DOMAIN 147 312
 FT DOMAIN 319 687
 FT DOMAIN 688 799
 FT DOMAIN 800 1184
 FT DOMAIN 1185 1573
 FT DOMAIN 1574 1968
 FT DOMAIN 1969 2370
 FT DOMAIN 2366 2402
 FT DISULFID 2370 2381
 FT DISULFID 2375 2390
 FT DISULFID 2392 2401
 FT CARBOHYD 109 109
 FT CARBOHYD 269 269
 FT CARBOHYD 735 735
 FT CARBOHYD 758 758
 FT CARBOHYD 833 833
 FT CARBOHYD 1154 1154
 FT CARBOHYD 1329 1329
 FT CARBOHYD 1448 1448
 FT CARBOHYD 1544 1544
 FT CARBOHYD 1596 1596
 FT CARBOHYD 1654 1654
 FT CARBOHYD 1843 1843
 FT CARBOHYD 1965 1965
 FT CARBOHYD 2122 2122
 FT CARBOHYD 2165 2165
 FT CARBOHYD 2178 2178
 FT CARBOHYD 2329 2329
 FT CARBOHYD 2359 2359
 FT CONFLICT 823 823
 FT CONFLICT 923 923
 FT CONFLICT 965 965
 FT CONFLICT 1241 1241
 SQ SEQUENCE 2476 AA; 270364 MW; 138690375A6548C CRC64;
 Query Match 27.5%; Score 107.5; DB 1; Length 2476;
 Best Local Similarity 33.3%; Pred. No. 0.00071;
 Matches 22; Conservative 9; Mismatches 18; Indels 17; Gaps 3;
 QY 11 KCPNSIEFRCDCRC-----QRFPCNVVVKPLCIKICAPGCVCRGLGYLRNKKVKCV 61
 DB 1455 KCPGSSYSTCANPCPATCLSLNPNYCPSTLP-----CAEGCEQKGHILSGTS-CV 1506
 QY 62 PRSKCG 67
 DB 1507 PLSQCG 1512
 RESULT 4
 ID AMCI_APIME STANDARD; PRT; 56 AA.
 AC P56682;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Chymotrypsin inhibitor (AMCI).
 OS Apis mellifera (Honeybee).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
 OC Aculeata; Apoidea; Apidae; Apis.
 OX NCBI_TaxID=7460;
 RN SEQUENCE, AND STRUCTURE BY NMR.
 RC TISSUE=Hemolymph;
 RX MEDLINE=99339935; PubMed=10411628;
 RA Bania J., Stachowiak D., Polanowski A.;
 RT "Primary structure and properties of the cathepsin G/chymotrypsin
 inhibitor from the larval hemolymph of Apis mellifera.";
 RL Eur. J. Biochem. 262:680-687(1999).
 CC -!- FUNCTION: CHYMOTRYPSIN AND CATHEPSIN G INHIBITOR.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 DR PDB; ICCV; 12-MAR-99.
 DR InterPro; IPR002919; TIL_Cysrich.
 DR Pfam; PF01826; TIL_1.
 KW Serine protease inhibitor; 3D-structure.
 FT DISULFID 3 36
 FT DISULFID 12 32
 FT DISULFID 16 28
 FT DISULFID 20 56
 FT DISULFID 38 50
 SQ SEQUENCE 56 AA; 5973 MW; 092B2815AE6B2B7F CRC64;
 Query Match 26.7%; Score 104.5; DB 1; Length 56;
 Best Local Similarity 36.8%; Pred. No. 7.6e-05;
 Matches 21; Conservative 7; Mismatches 26; Indels 3; Gaps 2;
 QY 11 KCPNSIEFRCDCRCQRFPCNVVVKP-LCIKICAPGCVCRGLGYLRNKKVKCVPRSKC 66
 DB 2 ECGPNEVFNTCGSACAPTCAQ--PKTRICTMQCRIGCCQCEGFLRNGEGACVLPENC 56
 RESULT 5
 ID VWF_CANFA STANDARD; PRT; 2813 AA.
 AC Q28295; Q28311; Q9TS14;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 18-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Von Willebrand factor precursor (VWF).
 GN F8VWF OR VWF.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN SEQUENCE FROM N.A.
 RA Stoy S.J., Shibuya H., Nonneman D.J., Holzhauser J., Mohammed I.H.,
 RA Johnson G.S.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Montgomery R.R., Fahs S., Montgomery M.W.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Venta P.J., Li J., Yuzbasiyan-Gurkan V., Brewer G.J., Schall W.D.;
 RT "Complete sequence of the structural gene for canine von Willebrand
 factor and identification of a mutation causing Scottish terrier von
 Willebrand's disease.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1234-1669 FROM N.A.
 RC TISSUE=Blood;
 RA Mancuso D.J., Christopherson P.A., Kroner P.A., Montgomery R.R.;
 RT "The canine von Willebrand factor gene: sequence and expression of
 a region encoding the glycoprotein Ib/IX binding domain.";
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: IMPORTANT IN THE MAINTENANCE OF HOMEOSTASIS. IT
 CC PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A
 CC NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; U85763; AAB96387.1; -
 EMBL; AE003475; AAR47683.1; -
 EMBL; AY010608; AAG35367.1; -
 EMBL; AY010609; AAG35368.1; -
 EMBL; AY010610; AAG35369.1; -
 EMBL; AY010611; AAG35370.1; -
 EMBL; AY010612; AAG35371.1; -
 EMBL; AY010613; AAG35372.1; -
 EMBL; AY010614; AAG35373.1; -
 EMBL; AY010615; AAG35374.1; -
 EMBL; AY010616; AAG35375.1; -
 EMBL; AY010617; AAG35376.1; -
 FlyBase; FBgn020509; ACP62F.
 InterPro; IPR002919; rTL_Cysrich.
 Pfam; PF01826; TIL; 1.
 Behavior; Signal.
 KW SIGNAL 1
 FT CHAIN 25 115 ACCESSORY GLAND PROTEIN ACP62F.
 SQ SEQUENCE 115 AA; 12570 MW; 4326A6F6C32291D CRC64;

Query Match 25 18; Score 98; DB 1; Length 115;
 Best Local Similarity 34.48; Pred. No. 0.00065;
 Matches 22; Conservative 9; Mismatches 31; Indels 2; Gaps 2;

QY 4 GGLGGRGKCPNEIFSRCDGRCPNVVPKPLCIKICAPGCVCLGYLNRKK-KVCVP 62
 DB 26 GMOGPKVDCGTANGTQTECPVACPCEYSGNGP-CVRKMGAPCVCKPGYVNERIPACVL 84

QY 63 RSKC 66
 DB 85 RSDC 88

RESULT 9

VWF_HUMAN STANDARD; PRT; 2813 AA.
 AC P04275;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Von Willebrand factor precursor (vWF).
 GN F8VWF OR VWF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90062044; PubMed=2584182;
 RA Mancuso D.J., Tuley E.A., Westfield L.A., Worrall N.K.,
 RA Shelton-Inloes B.B., Sorace J.M., Alevis Y.G., Sadler J.E.;
 RT "Structure of the gene for human von Willebrand factor.";
 RL J. Biol. Chem. 264:19514-19527(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87016349; PubMed=3489923;
 RA Bonthron D., Orr E.C., Mitsuoka L.M., Ginsburg D., Handin R.I.,
 RA Orkin S.H.;
 RT "Nucleotide sequence of pre-pro-von Willebrand factor cDNA.";
 RL Nucleic Acids Res. 14:7125-7128(1986).
 RN [3]
 RP SEQUENCE OF 1-120 FROM N.A., AND SEQUENCE OF 23-56.
 RC TISSUE-Umbilical vein endothelial cells;
 RX MEDLINE=87213253; PubMed=3495266;
 RA Shelton-Inloes B.B., Broze G.J. Jr., Miletich J.P., Sadler J.E.;

RT "Evolution of human von Willebrand factor: cDNA sequence
 RT polymorphisms, repeated domains, and relationship to von Willebrand
 RT antigen II.";
 RL Biochem. Biophys. Res. Commun. 144:657-665(1987).
 RN [4]
 RP SEQUENCE OF 1-1400 FROM N.A.
 RX MEDLINE=87004550; PubMed=3019665;
 RA Verweij C.L., Diergaarde P.J., Hart M., Pannekoek H.;
 RT "Full-length von Willebrand factor (vWF) cDNA encodes a highly
 RT repetitive protein considerably larger than the mature vWF subunit.";
 RL EMBO J. 5:1839-1847(1986).
 RN [5]
 RP ERRATUM.
 RA Verweij C.L., Diergaarde P.J., Hart M., Pannekoek H.;
 RL EMBO J. 5:3074-3074(1986).
 RN [6]
 RP SEQUENCE OF 764-2813.
 RX MEDLINE=86269895; PubMed=3524673;
 RA Titani K., Kumar S., Takio K., Ericsson L.H., Wade R.D., Ashida K.,
 RA Walsh K.A., Chopek M.W., Sadler J.E., Fujikawa K.;
 RT "Amino acid sequence of human von Willebrand factor.";
 RL Biochemistry 25:3171-3184(1986).
 RN [7]
 RP SEQUENCE OF 781-1424 FROM N.A.
 RX MEDLINE=86269894; PubMed=3488076;
 RA Shelton-Inloes B.B., Titani K., Sadler J.E.;
 RT "cDNA sequences for human von Willebrand factor reveal five types of
 RT repeated domains and five possible protein sequence polymorphisms.";
 RL Biochemistry 25:3164-3171(1986).
 RN [8]
 RP SEQUENCE OF 764-873 AND 1289-2813 FROM N.A.
 RX MEDLINE=86016708; PubMed=2864688;
 RA Sadler J.E., Shelton-Inloes B.B., Sorace J.M., Harlan J.M.,
 RA Titani K., Davie E.W.;
 RT "Cloning and characterization of two cDNAs coding for human von
 RT Willebrand factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:6394-6398(1985).
 RN [9]
 RP SEQUENCE OF 990-1947 FROM N.A.
 RX MEDLINE=91105089; PubMed=1988024;
 RA Mancuso D.J., Tuley E.A., Westfield L.A., Lester-Mancuso T.L.,
 RA Le Beau M.M., Sorace J.M., Sadler J.E.;
 RT "Human von Willebrand factor gene and pseudogene: structural analysis
 RT and differentiation by polymerase chain reaction.";
 RL Biochemistry 30:253-269(1991).
 RN [10]
 RP SEQUENCE OF 2731-2813 FROM N.A.
 RX MEDLINE=85269603; PubMed=3875078;
 RA Verweij C.L., de Vries C.J.M., Distel B., van Zonneveld A.-J.,
 RA Geurts van Kessel A., van Mourik J.A., Pannekoek H.;
 RT "Construction of cDNA coding for human von Willebrand factor using
 RT antibody probes for colony-screening and mapping of the chromosomal
 RT gene.";
 RL Nucleic Acids Res. 13:4699-4717(1985).
 RN [11]
 RP SEQUENCE OF 1-177 FROM N.A.
 RX MEDLINE=88111704; PubMed=2828057;
 RA Bonthron D., Orkin S.H.;
 RT "The human von Willebrand factor gene. Structure of the 5' region.";
 RL Eur. J. Biochem. 171:51-57(1988).
 RN [12]
 RP SEQUENCE OF 2621-2813 FROM N.A.
 RX MEDLINE=85244588; PubMed=3874428;
 RA Ginsburg D., Handin R.I., Bonthron D.T., Donlon T.A., Bruns G.A.P.,
 RA Latt S.A., Orkin S.H.;
 RT "Human von Willebrand factor (vWF): isolation of complementary DNA
 RT (cDNA) clones and chromosomal localization.";
 RL Science 228:1401-1406(1985).
 RN [13]
 RP SEQUENCE OF 2731-2813 FROM N.A.
 RX MEDLINE=85201687; PubMed=3873280;
 RA Lynch D.C., Zimmerman T.S., Collins C.J., Brown M., Morin M.J.,
 RA Ling E.H., Livingston D.M.;

RT "Molecular cloning of cDNA for human von Willebrand factor:
 RT authentication by a new method."; [14]
 RN
 RA Lynch D.C.; [15]
 RA Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE OF 2731-2813 FROM N.A.
 RN MEDLINE=87260814; PubMed=3496594;
 RA Collins C.J., Underdahl J.P., Levene R.B., Ravera C.P.,
 RA Morin M.J., Dombalagian M.J., Ricca G., Livingston D.M.,
 RA Lynch D.C.;
 RA "Molecular cloning of the human gene for von Willebrand factor and
 RT identification of the transcription initiation site."; [16]
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4393-4397(1987).
 RN
 RP DISULFIDE BONDS.
 RA MEDLINE=88163465; PubMed=3502076;
 RA Marti T., Rosselet S.J., Titani K., Walsh K.A.;
 RT "Identification of disulfide-bridged substructures within human von
 RT Willebrand factor."; [17]
 RL Biochemistry 26:8099-8109(1987).
 RN
 RP STRUCTURE OF CARBOHYDRATES.
 RA MEDLINE=86274702; PubMed=3089784;
 RA Samor B., Michalski J.C., Debray H., Mazurier C., Goudemand M.,
 RA van Halbeek H., Vliegthart J.F.G., Montreuil J.;
 RT "Primary structure of a new tetraantennary glycan of the N-
 RT acetylglucosaminic type isolated from human factor VIII/von
 RT Willebrand factor."; [18]
 RL Eur. J. Biochem. 158:295-298(1986).
 RN
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 1261-1468.
 RX MEDLINE=98221174; PubMed=9553097;
 RA Emsley J., Cruz M., Handin R., Liddington R.;
 RT "Crystal structure of the von Willebrand factor A1 domain and
 RT implications for the binding of platelet glycoprotein Ib."; [19]
 RL J. Biol. Chem. 273:10396-10401(1998).
 RN
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1685-1873.
 RX MEDLINE=97472999; PubMed=9331419;
 RA Huizinga E.G., Martijn van der Plas R., Kroon J., Sixma J.J., Gros P.;
 RT "Crystal structure of the A3 domain of human von Willebrand factor:
 RT implications for collagen binding."; [20]
 RL Structure 5:1147-1156(1997).
 RN
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1686-1872.
 RX MEDLINE=97460108; PubMed=9312128;
 RA Blenkowska J., Cruz M., Atiemo A., Handin R., Liddington R.;
 RT "The von Willebrand factor A3 domain does not contain a metal ion-
 RT dependent adhesion site motif."; [21]
 RL J. Biol. Chem. 272:25162-25167(1997).
 RN
 RP VARIANTS TRP-1597 AND ASP-1607.
 RX MEDLINE=89264495; PubMed=2786201;
 RA Ginsburg D., Konkle B.A., Gill J.C., Montgomery R.R.,
 RA Bockenstedt P.L., Johnson T.A., Yang A.Y.;
 RT "Molecular basis of human von Willebrand disease: analysis of
 RT platelet von Willebrand factor mRNA."; [22]
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3723-3727(1989).
 RN
 RP VARIANT THR-1628.
 RX MEDLINE=91196734; PubMed=1673047;
 RA Iannuzzi M.C., Hidak N., Boehnke M., Bruck M.E., Hanna W.T.,
 RA Collins F.S., Ginsburg D.;
 RT "Analysis of the relationship of von Willebrand disease (vwd) and
 RT hereditary hemorrhagic telangiectasia and identification of a
 RT potential type IIA vwd mutation (Ile865 to Thr)."; [23]
 RL Am. J. Hum. Genet. 48:757-763(1991).
 RN
 RP VARIANTS NORMANDY-2 AND NORMANDY-3.
 RX MEDLINE=92001464; PubMed=1832934;

RA Gaucher C., Mercier B., Jorieux S., Oufkir D., Mazurier C.;
 RT "Identification of two point mutations in the von Willebrand factor
 RT gene of three families with the 'Normandy' variant of von Willebrand
 RT disease."; [24]
 RL Br. J. Haematol. 78:506-514(1991).
 RN
 RP VARIANT CYS-1308.
 RX MEDLINE=92104315; PubMed=1761120;
 RA Donner M., Andersson A.-M., Kristoffersson A.-C., Nilsson I.M.,
 RA Dahlback B., Holmberg L.;
 RT "An Arg545-->Cys545 substitution mutation of the von Willebrand
 RT factor in type IIB von Willebrand's disease."; [25]
 RL Eur. J. Haematol. 47:342-345(1991).
 RN
 RP VARIANTS TRP-1306; CYS-1308 AND PRO-1613.
 RX MEDLINE=91185601; PubMed=2010538;
 RA Randi A.M., Rabinowitz I., Mancuso D.J., Mannucci P.M., Sadler J.E.;
 RT "Molecular basis of von Willebrand disease type IIB. Candidate
 RT mutations cluster in one disulfide loop between proposed platelet
 RT glycoprotein Ib binding sequences."; [26]
 RL J. Clin. Invest. 87:1220-1226(1991).
 RN
 RP VARIANTS TRP-1306; CYS-1308; MET-1316; GLN-1341 AND HIS-1399.
 RX MEDLINE=91185602; PubMed=1672694;
 RN
 Query Match 25.1%; Score 98; DB 1; Length 2813;
 Best Local Similarity 26.0%; Pred. No. 0.0079;
 Matches 19; Conservative 13; Mismatches 27; Indels 14; Gaps 2;
 QY 8 GRG-----KCPSEIFSRCDGRQCFQPNV-VKPLCKIKICAPGCVRLGYL 53
 DB 635 GRGVAVREPGRCGLNCPKGVYLCQGTPTCNLTCSRSLSPDEECNEACLEGCFPPGLY 694
 QY 54 RNKKKVCVPRSKC 66
 DB 695 MDERGDCVPEKQC 707
 RESULT 10
 VWF_PIG
 ID VWF_PIG STANDARD; PRT; 2482 AA.
 AC Q28833;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Von Willebrand factor precursor (VWF) (Fragment).
 GN F8VWF OR VWF.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seaman W.T., Read M.S., Bellinger D.A., Nichols T.C.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 397-553 FROM N.A.
 RX MEDLINE=93356762; PubMed=8352759;
 RA Lavergne J.M., Piao Y.C., Ferreira V., Kerbirou-Nabias D.,
 RA Bahnak B.R., Meyer D.;
 RT "Primary structure of the factor VIII binding domain of human, porcine
 RT and rabbit von Willebrand factor."; [3]
 RL Biochem. Biophys. Res. Commun. 194:1019-1024(1993).
 CC
 CC !- FUNCTION: IMPORTANT IN THE MAINTENANCE OF HOMEOSTASIS, IT
 CC PARTICIPATES IN PLATELET-VESSEL WALL INTERACTIONS BY FORMING A
 CC NONCOVALENT COMPLEX WITH COAGULATION FACTOR VIII AT THE SITE OF
 CC VASCULAR INJURY (BY SIMILARITY).
 CC !- SUBUNIT: MULTIMERIC (BY SIMILARITY).
 CC !- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
 CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
 CC !- SIMILARITY: CONTAINS 3 VWF DOMAINS.
 CC !- SIMILARITY: CONTAINS 3 VWF DOMAINS.
 CC !- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

CC CC -!- SIMILARITY: SOME, TO SILKWORM HEMOCYTIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF052036; AAC06229.1; -;
 CC EMBL: S64541; AAB27829.2; -;
 CC HSSP: P04275; IATZ.
 CC InterPro: IPR000359; Cys_knot.
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR001928; Endothln_tox.
 CC InterPro: IPR000864; Potato_inhibit.
 CC InterPro: IPR002919; TIL_Cysrich.
 CC InterPro: IPR002035; VWF_A.
 CC InterPro: IPR001007; VWF_C.
 CC InterPro: IPR001846; VWF_D.
 CC Pfam: PF00007; Cys_knot; 1.
 CC Pfam: PF00092; wa; 3.
 CC Pfam: PF00093; wvc; 3.
 CC Pfam: PF00094; wvd; 3.
 CC Pfam: PF01826; TIL; 3.
 CC PRINTS: PR00365; ENDOTHELIN.
 CC PRINTS: PR00292; POTATOINHTR.
 CC PRINTS: PR00453; VWFADOMAIN.
 CC SMART: SM00041; CT; 1.
 CC SMART: SM00181; EGF; 1.
 CC SMART: SM00327; VWA; 3.
 CC SMART: SM00214; VWC; 5.
 CC SMART: SM00216; VWD; 3.
 CC PROSITE: PS02334; VWEA; 3.
 CC PROSITE: PS01208; VWFC; 3.
 CC PROSITE: PS01185; CTCK_1; 1.
 CC PROSITE: PS01225; CTCK_2; 1.
 CC Blood coagulation; Platelet; Glycoprotein; Extracellular matrix;
 CC Plasma; Hemostasis; Repeat; Cell adhesion.
 CC NON_TER 1 1
 CC PROPEP <1 437 BY SIMILARITY.
 CC CHAIN 438 2482 VON WILLEBRAND FACTOR.
 CC DOMAIN 62 215 WFED 2.
 CC DOMAIN 438 461 AMINO-TERMINAL.
 CC DOMAIN 462 507 E1.
 CC DOMAIN 500 527 CX.
 CC DOMAIN 541 687 WFED 3.
 CC DOMAIN 947 1127 WVEA 1.
 CC DOMAIN 1167 1334 WVEA 2.
 CC DOMAIN 1360 1540 WVEA 3.
 CC DOMAIN 1619 1771 WFED 4.
 CC DOMAIN 1885 1930 E2.
 CC DOMAIN 1924 1997 WFEC 1.
 CC DOMAIN 2098 2164 WFEC 2.
 CC DOMAIN 2249 2319 WFEC 3.
 CC DOMAIN 2393 2481 CTCK.
 CC SITE 2176 2178 CELL ATTACHMENT SITE (POTENTIAL).
 CC DISULFID 441 482 BY SIMILARITY.
 CC DISULFID 450 478 BY SIMILARITY.
 CC DISULFID 484 495 BY SIMILARITY.
 CC DISULFID 541 670 BY SIMILARITY.
 CC DISULFID 563 705 BY SIMILARITY.
 CC DISULFID 572 667 BY SIMILARITY.
 CC DISULFID 588 595 BY SIMILARITY.
 CC DISULFID 734 758 BY SIMILARITY.
 CC DISULFID 745 785 BY SIMILARITY.
 CC DISULFID 763 765 BY SIMILARITY.
 CC DISULFID 827 839 BY SIMILARITY.
 CC DISULFID 823 843 BY SIMILARITY.
 CC DISULFID 800 804 BY SIMILARITY.
 CC DISULFID 870 873 BY SIMILARITY.
 CC DISULFID 908 911 BY SIMILARITY.

FT DISULFID 942 1128 BY SIMILARITY.
 FT DISULFID 1338 1339 BY SIMILARITY.
 FT DISULFID 1355 1541 BY SIMILARITY.
 FT DISULFID 1548 1573 BY SIMILARITY.
 FT DISULFID 1568 1609 OR 1611 (BY SIMILARITY).
 FT DISULFID 1641 1792 BY SIMILARITY.
 FT DISULFID 1619 1754 BY SIMILARITY.
 FT DISULFID 1596 1757 BY SIMILARITY.
 FT DISULFID 1662 1670 BY SIMILARITY.
 FT DISULFID 2393 2443 BY SIMILARITY.
 FT DISULFID 2408 2457 BY SIMILARITY.
 FT DISULFID 2419 2473 BY SIMILARITY.
 FT DISULFID 2423 2475 BY SIMILARITY.
 FT DISULFID ? 2480 BY SIMILARITY.
 FT CARBOHYD 340 340 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 905 905 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1184 1184 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1640 1640 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1892 1892 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1959 1959 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2026 2026 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2069 2069 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2215 2215 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2254 2254 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 2459 2459 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT SEQUENCE 2482 AA; 272394 MW; D499B7DDFBCEADD CRC64;
 SQ
 Query Match 23.4%; Score 91.5; DB 1; Length 2482;
 Best Local Similarity 28.8%; Pred. No. 0.035;
 Matches 19; Conservative 11; Mismatches 31; Indels 5; Gaps 2;
 QY 2 GFGGLGGRGKCFSPNEIFSCDGRQCFPNV-VPKPLKIKICAPGCVRLGYLRNKKVC 60
 Db 320 GFCAL----SCPPGQVYLQCGTFCNLTCRSLSPDECAEDCLGFCFPPGLYLDGSGDC 375
 QY 61 VPRSKC 66
 Db 376 VPKAQ 381
 RESULT 11
 HMCT_BOMMO STANDARD; PRT; 3133 AA.
 ID HMCT_BOMMO STANDARD; PRT; 3133 AA.
 AC P98092;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hemocytin precursor (Humoral lectin).
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 OC Ditrysia; Bombycoidea; Bombycidae; Bombyx.
 OX NCBI_TaxID=7091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FUYOU X Tokai; TISSUE=Hemocyte;
 RX MEDLINE=95178544; PubMed=7873598;
 RA Kotani E., Yamakawa M., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,
 RA Matsubara F., Tanai K., Kadono-Okuda K., Kato Y., Mori H.;
 RT "Cloning and expression of the gene of hemocytin, an insect humoral
 RT lectin which is homologous with the mammalian von Willebrand
 RT factor.";
 RL Blochim. Biophys. Acta 1260:245-258(1995).
 RN [2]
 RP SEQUENCE OF 2221-3133 FROM N.A.
 RA Kotani E., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,
 RA Matsubara F., Yamakawa M.;
 RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ADHESIVE PROTEIN AND RELATES TO HEMOSTASIS OR
 CC ENCAPSULATION OF FOREIGN SUBSTANCES FOR SELF-DEFENSE.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN HEMOCYTES DURING LARVAL-PUPAL
 CC METAMORPHOSIS.

CC CC -(- INDUCTION: HEMAGGLUTINATION ACTIVITY IS INCREASED BY BACTERIAL
OR VIRAL INFECTION AND INHIBITED BY D-MANNOSE, N-ACETYL-D-
GALACTOSAMINE AND D-MALTOSE.
CC CC -(- PPM: MAY BE CONVERTED INTO THE 260 kda MAJURE HEMOCYTIN BY
PROTEOLYSIS.
CC CC -(- SIMILARITY: TO MAMMALIAN VON WILLEBRAND FACTOR.
CC CC -(- SIMILARITY: THE N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT
OF HUMAN MUCIN 2.
CC CC -(- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.
CC CC -(- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC CC -(- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC CC -----
CC EMBL; D29738; BAA06160.1; -;
DR EMBL; D14035; BAA03124.1; -;
DR HSSP; P12259; 1C2T.
DR InterPro; IPR000359; Cys_Knot.
DR InterPro; IPR000421; FA58_C.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR001007; VWF_C.
DR InterPro; IPR001846; VWF_D.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00094; vwd_3.
DR Pfam; PF00754; F5_F8_type_C; 2.
DR Pfam; PF01826; TIL; 6.
DR SMART; SM00041; CT; 1.
DR SMART; SM00231; FA58C; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00192; LDLA; 1.
DR SMART; SM00214; VWC; 3.
DR SMART; SM00216; VWD; 3.
DR PROSITE; PS01185; znf_C2HC; 1.
DR PROSITE; PS01208; VWF_C; FALSE_NEG.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS01285; FA58C_1; 2.
DR PROSITE; PS01286; FA58C_2; 2.
KW Lentin; Glycoprotein; Signal; Repeat; Cell adhesion.
FT SIGNAL 1 ?
FT CHAIN ? 3133
FT HEMIN ? 131
FT DOMAIN 29 131
FT DOMAIN 153 240
FT DOMAIN 248 613
FT DOMAIN 940 1095
FT DOMAIN 1116 1254
FT DOMAIN 1283 1356
FT DOMAIN 1620 1951
FT DOMAIN 1952 2315
FT DOMAIN 2330 2321
FT DOMAIN 2335 2361
FT DOMAIN 2435 2469
FT DOMAIN 2553 2622
FT DOMAIN 2842 2907
FT DOMAIN 2971 3076
FT DOMAIN 895 914
FT DOMAIN 1267 1270
FT DOMAIN 1425 1428
FT DOMAIN 1447 1450
FT DOMAIN 1474 1479
FT DOMAIN 2148 2153
FT DOMAIN 2156 2159
FT DOMAIN 2341 2344
FT DISULFID 940 1095
FT DISULFID 1116 1254
BY SIMILARITY.

FT DISULFID 2981 3040 BY SIMILARITY.
FT DISULFID 2991 3054 BY SIMILARITY.
FT DISULFID 3004 3070 BY SIMILARITY.
FT DISULFID 3020 3072 BY SIMILARITY.
FT DISULFID ? 3075 BY SIMILARITY.
FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1170 1170 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1387 1387 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1622 1622 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1727 1727 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1847 1847 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1975 1975 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1985 1985 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2093 2093 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2113 2113 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2161 2161 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2276 2276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2451 2451 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2647 2647 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2654 2654 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2663 2663 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2794 2794 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2810 2810 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2865 2865 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2929 2929 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2964 2964 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3028 3028 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 1288 1288 R -> G.
FT VARIANT 1305 1305 T -> S.
SQ SEQUENCE 3133 AA; 343350 MW; E5210D5D14A7B2B2 CRC64;
Query Match 23.1%; Score 90.5; DB 1; Length 3133;
Best Local Similarity 27.5%; Pred. No. 0.053;
Matches 19; Conservative 11; Mismatches 26; Indels 13; Gaps 3;
QY 11 KCPSEIFSRCDGRCORFCPN---VVPKPLGI-KICAPGV-----CRLGYLRNKK 57
Db 769 RCPGGEVQACAYKCDRLCDHFKTLIAKRCISEMCDGCVDSVASNGCESSWRDE 828
QY 58 KVCVPRSKC 66
Db 829 RTCVPVKDC 837
RESULT 12
BAR3_CHITE ID BAR3_CHITE STANDARD; PRT; 1700 AA.
AC Q03376;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Balbiani ring protein 3 precursor.
GN BR3.
OS Chironomus tentans (Midge).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Chironomoidea; Chironomidae; Chironominae; Chironomus.
RN NCBI_TaxID=7153;
[1]
SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=90172404; PubMed=1689777;
RA Paulsson G., Lendahl U., Galli J., Ericsson C., Wieslander L.;
RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
repetitive structure split by many introns.";
RL J. Mol. Biol. 211:331-349(1990).
CC -(- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.

```

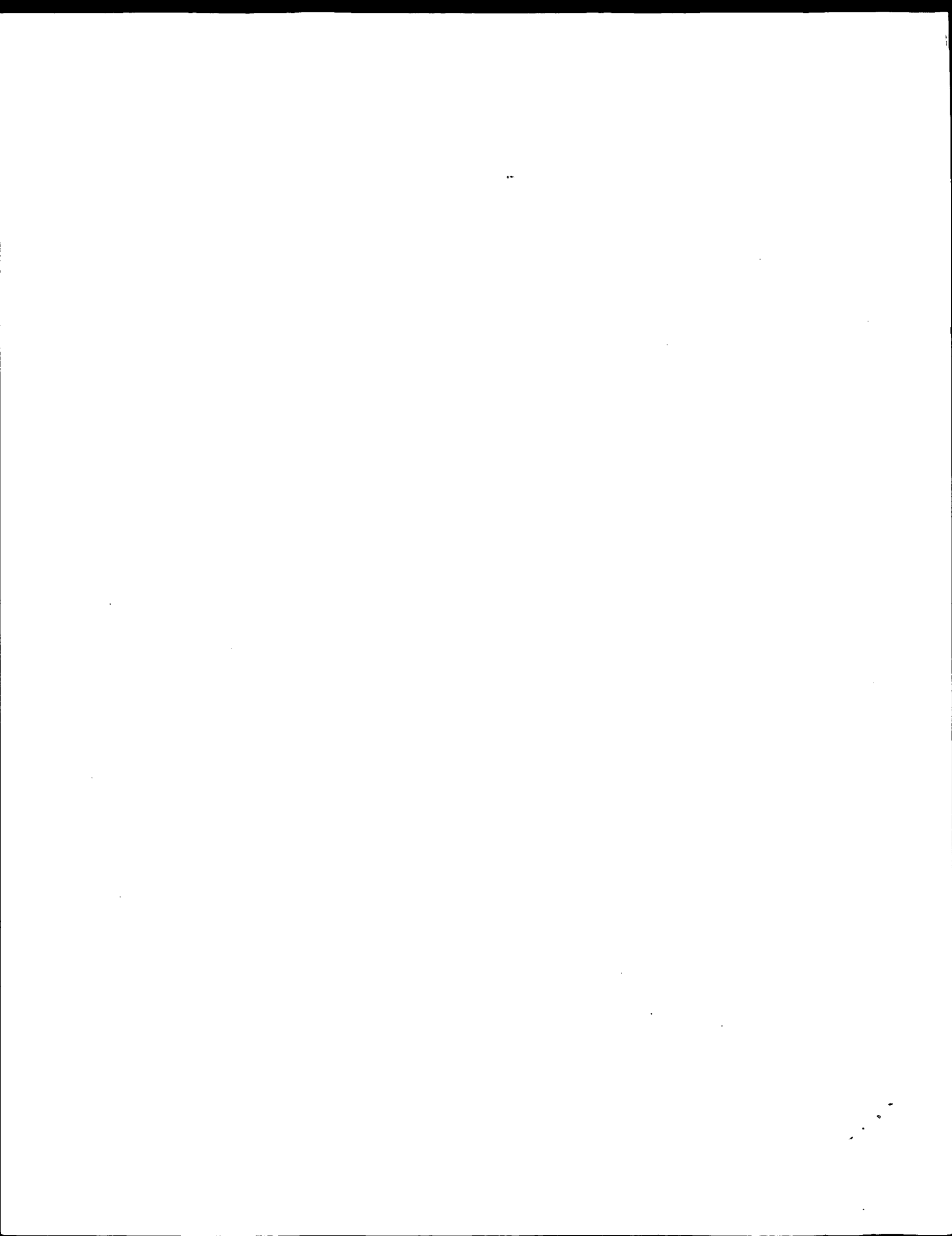
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: SALIVARY GLAND.
CC -!- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52263; CAA36506.1; -
DR PIR: S08167; S08167.
DR HSSP: P15358; 1SKZ.
DR InterPro: IPR004153; CXCXC_repeat.
DR Pfam: PF03128; CXCXC; 71.
KW Repeat; Signal.
FT SIGNAL 1 20
FT CHAIN 21 1700 BALBIANI RING PROTEIN 3.
FT SEQUENCE 1700 AA; 186145 MW; 34202828521B0815 CRC64;
SQ
Query Match 22.08; Score 86; DB 1; Length 1700;
Best Local Similarity 29.1%; Pred. No. 0.099;
Matches 23; Conservative 9; Mismatches 21; Indels 26; Gaps 6;
QY 12 CPSNEIFSRCDGRCQRCPCNVVVKP-----LCIKICAPGCV---CRILGYLRNK 56
DB 1234 CPGNOIW--CDTCRCVCPKNMERPADNCKTKWNDEMCCQCKPQSGGCK-GVMKN 1290
QY 57 KKVC---VPRSK-----CG 67
DB 1291 ANTSCCEPADKAPASCG 1309
RESULT 13
ID ICEL_ASCSU STANDARD; PRT; 63 AA.
AC P07851;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DE Chymotrypsin/elastase isoform 1 (C/E-1 inhibitor).
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE.
RX MEDLINE=84255715; PubMed=6564898;
RA Babin D.R., Peanasky R.J., Goos S.M.;
RT "The isoforms of chymotrypsin/elastase from Ascaris
RL lumbricoides: the primary structure.";
RL Arch. Biochem. Biophys. 232:143-161(1984).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF COMPLEX WITH ELASTASE.
RX MEDLINE=95006335; PubMed=7922044;
RA Huang K., Strynadka N.C., Bernard V.D., Peanasky R.J., James M.N.;
RT "The molecular structure of the complex of Ascaris
RL chymotrypsin/elastase inhibitor with porcine elastase.";
RL Structure 2:679-689(1994).
CC -!- FUNCTION: DEFEND THE ORGANISM AGAINST THE HOST'S PROTEINASES.
CC -!- SIMILARITY: BELONGS TO THE ASCARIS TRYPSIN INHIBITOR FAMILY.
DR PIR: S07127; S07127.
DR PDB: IEAI; 05-APR-99.
DR InterPro: IPR002919; TIL_Cysrich.
DR Pfam: PF01826; TIL; 1.
KW Serine protease inhibitor; 3D-structure.
FT DISULFID 5 38
FT DISULFID 14 33
FT DISULFID 17 29
FT DISULFID 21 60
FT DISULFID 40 54

```

```

FT ACT_SITE 31 32 REACTIVE BOND.
SQ SEQUENCE 63 AA; 6862 MW; 5DC10DE75B375F16 CRC64;
Query Match 20.6%; Score 80.5; DB 1; Length 63;
Best Local Similarity 30.3%; Pred. No. 0.029;
Matches 20; Conservative 10; Mismatches 23; Indels 13; Gaps 4;
QY 8 GRCKCPSNEIFSRCDG---RC-----QRFCPCNVVVKPLCIKICAPGCVCRILGYLRNK 60
DB 1 GQSCGCPNEVWTECTCEMKCGPDENTPCLMCRRPSC--ECSPG-----RGMRRNTDNGKC 54
QY 61 VPRSK 66
DB 55 IPASQC 60
RESULT 14
ID YOJ2_CAEEL STANDARD; PRT; 99 AA.
AC P34625;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE Hypothetical 10.6 kDa protein ZK353.2 in chromosome III.
GN ZK353.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Eavell A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins J., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L15313; AAA28199.1; -
DR WormPep: ZK353.2; CE00386.
KW Hypothetical protein.
SQ SEQUENCE 99 AA; 10561 MW; 862C659838E47E5F CRC64;
Query Match 20.6%; Score 80.5; DB 1; Length 99;
Best Local Similarity 28.1%; Pred. No. 0.041;
Matches 18; Conservative 10; Mismatches 17; Indels 19; Gaps 2;
QY 1 GGFGLGGRGKSPNEIFSRCDGRCQRCPCN-----VPRKPLCIKICA 43
DB 31 GGYGGYGGYGGCGADNVFYR--WRCCDYSPECCIQLETFWVFLVFIIGFFVCLACL 88
QY 44 PGCV 47
DB 89 AGCV 92

```

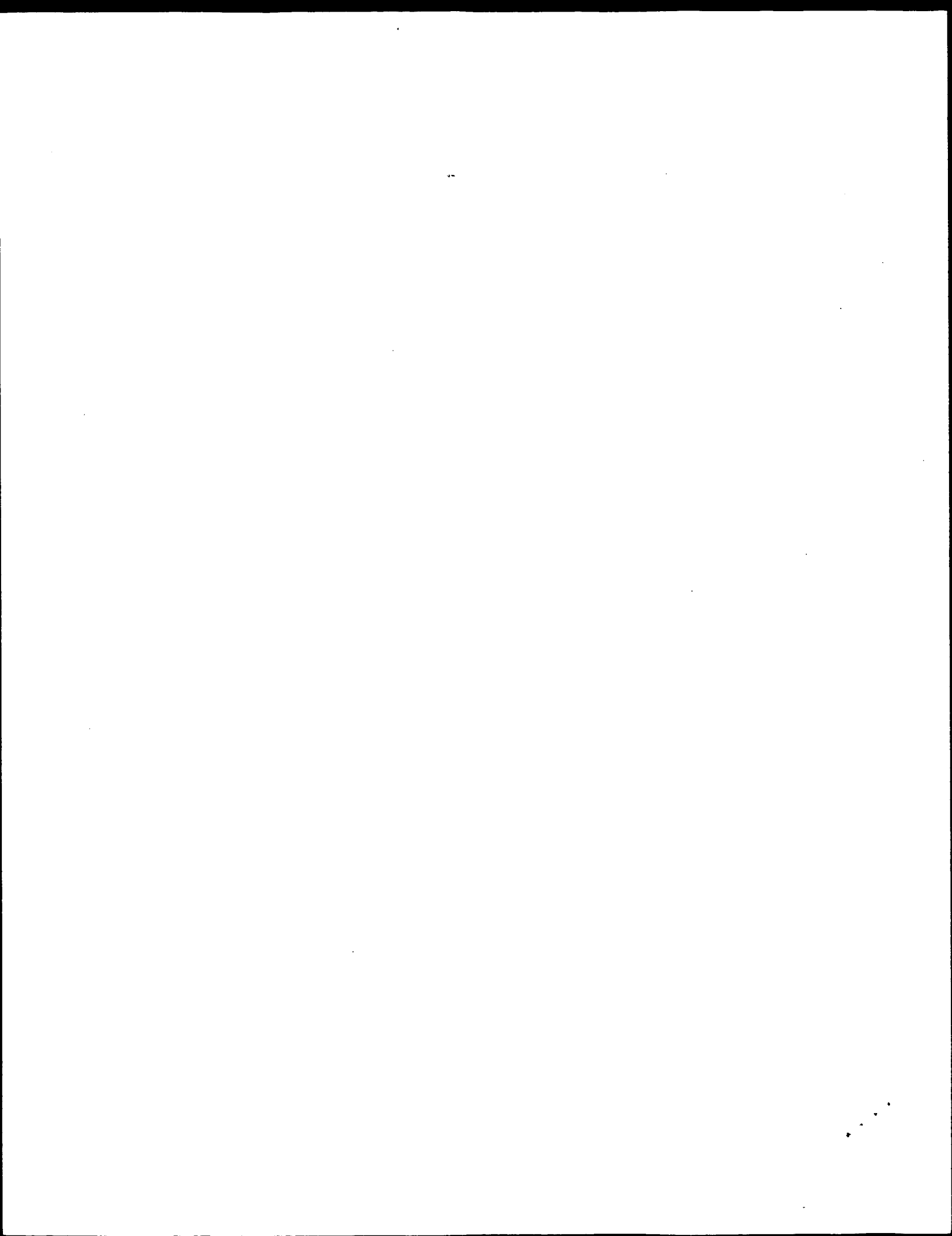



RT Dom.

RT Domain Structure." ;

DR	PROSITE; PS00740; MAM_1; UNKNOWN_1.
DR	PROSITE; PS50060; MAM_2; 3.
*KW	Glycoprotein.
SQ	SEQUENCE 2689 AA; 291277 MW; 2253CC9E2D2D7C45F CRC64;
Query Match	27.9%; Score 109; DB 4; Length 2689;
Best Local Similarity	35.3%; Pred. No. 1.4e-05;
Matches	24; Conservative 11; Mismatches 13; Indels 20; Gaps 4;
Qy	11 KCPNSEIFSRC-----DGRCQFCPNVVPKPLCIKICAPGCVCRGLYLNRNKKV 59 ::: :: :: ::
Dd	2210 ECPAYSYTNCLPSCSPSCWDLGRCE---GAKVP-----SACAEGCICPGYVLSEDK- 2260 ::
Qy	60 CVPRSKCG 67
Dd	2261 CVPRSQCG 2268
RESULT 15	
Q9BZ83	PRELIMINARY; PRT; 2721 AA.
ID	O9BZ83
AC	O9BZ83;
DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Zonadhesin variant 6.
GN	Homo sapiens (Human).
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	TISSUE=TESTIS;
RC	Cheung T.L., Wassler M.J., Cornwall G.A., Hardy D.M.;
RT	"Multiple Intra-species Variants of Human Zonadhesin,"
RL	Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
CC	- - SIMILARITY: CONTAINS 3 MAM DOMAINS.
DR	EMBL; AF323980; AAK01436.1; "
DR	HSSP; P56682; ICCV.
DR	InterPro; IPR0003561; EGF-like.
DR	InterPro; IPR000998; MAM_domain.
DR	InterPro; IPR002965; P_rich_extensio.
DR	InterPro; IPR003328; TILA_Cysrich.
DR	InterPro; IPR002919; TIL_Cysrich.
DR	InterPro; IPR001846; VWF_D.
DR	Pfam; PF00008; EGF; 1.
DR	Pfam; PF00629; MAM; 3.
DR	Pfam; PF01826; TIL; 4.
DR	Pfam; PF02345; Tila; 4.
DR	Pfam; PF00094; vwd; 4.
DR	PRINTS; PR01217; PRICHEXTENS.
DR	SMART; SM00001; EGF_Like; 1.
DR	SMART; SM00137; MAM; 3.
DR	SMART; SM00216; VWD; 4.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR	PROSITE; PS01186; EGF_2; UNKNOWN_4.
DR	PROSITE; PS00740; MAM_1; 1.
DR	PROSITE; PS50060; MAM_2; 3.
KW	Glycoprotein.
SQ	SEQUENCE 2721 AA; 295506 MW; 142E9CBF3D40EC3 CRC64;
Query Match	27.9%; Score 109; DB 4; Length 2721;
Best Local Similarity	35.3%; Pred. No. 1.4e-05;
Matches	24; Conservative 11; Mismatches 13; Indels 20; Gaps 4;
Qy	11 KCPNSEIFSRC-----DGRCQFCPNVVPKPLCIKICAPGCVCRGLYLNRNKKV 59 ::: :: :: ::
Dd	2210 ECPAYSYNTNLPCSPSCWDLGRCE---GAKVP-----SACAEGCICPGYVLSEDK- 2260 ::

Search completed: February 26, 2003, 15:05:43
Job time : 31 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:03:23 ; Search time 36 Seconds
(without alignments)
247.994 Million cell updates/sec

Title: US-09-506-978-1
Perfect score: 391
Sequence: 1 GGFGGLGGKCPSEIFSR.....CRLGLRNKKKVCVPSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	391	100.0	67	ABB08330	Bee venom protein
2	391	100.0	69	ABB08331	Bee venom protein
3	391	100.0	71	ABB08332	Bee venom protein
4	391	100.0	73	ABB08333	Bee venom protein
5	391	100.0	92	AA199209	Amino acid sequenc
6	108	27.6	3843	ABB71529	Drosophila melanog
7	107.5	27.5	2476	AAW67738	Pig p105 zona pell
8	106.5	27.2	2594	AAW14748	IgG-Fc binding pro
9	106.5	27.2	2957	ABG22214	Novel human diagno
10	106.5	27.2	5405	AAW14749	IgG-Fc binding pro

11	101.5	26.0	7337	22	ABG22216	Novel human diagno
12	100.5	25.7	735	22	ABG22215	Novel human diagno
13	100	25.6	84	20	AA130432	Mature nematode ex
14	100	25.6	84	21	AA115317	A. caninum nematod
15	100	25.6	91	17	AA191701	AcanaPc2. Ancylos
16	100	25.6	91	20	AA130393	Nematode extracted
17	100	25.6	91	20	AA130454	Nematode extracted
18	100	25.6	91	21	AA115346	A. caninum nematod
19	98.5	25.2	508	22	ABG22213	Novel human diagno
20	98	25.1	115	20	AA122170	Drosophila Acp62F
21	98	25.1	115	22	ABB57966	Drosophila melanog
22	98	25.1	741	17	AA16462	Human von Willebra
23	98	25.1	2813	7	AA160053	Sequence of von Wi
24	98	25.1	2813	7	AA160462	Sequence of human
25	98	25.1	2813	23	AAU75317	Human von Willebra
26	98	25.1	2814	22	AAU29530	Novel human secret
27	96	24.6	164	22	AAE04265	Human gene 8 encod
28	96	24.6	164	22	AAE01850	Human gene 8 encod
29	96	24.6	189	22	AAE04268	Human gene 8 encod
30	96	24.6	191	22	AAE04266	Human gene 8 encod
31	95.5	24.4	149	22	AAE04263	Human gene 8 encod
32	92.5	23.7	108	22	ABB63313	Drosophila melanog
33	92	23.5	701	22	ABB66725	Drosophila melanog
34	91	23.3	82	20	AA130399	Nematode extracted
35	91	23.3	82	20	AA130422	Mature nematode ex
36	91	23.3	82	21	AA15293	A. ceylanicum nema
37	91	23.3	82	21	AA15307	A. ceylanicum nema
38	91	23.3	171	20	AA130435	Mature nematode ex
39	91	23.3	190	17	AA191710	AcenAP4. Ancylost
40	91	23.3	190	20	AA130384	Nematode extracted
41	88	22.5	2009	22	ABB64069	Drosophila melanog
42	87.5	22.4	178	22	ABB61389	Drosophila melanog
43	87.5	22.4	1438	22	ABG03631	Novel human diagno
44	87.5	22.4	1438	22	ABG28308	Novel human diagno
45	87.5	22.4	1439	22	ABG14849	Novel human diagno

ALIGNMENTS

RESULT 1

ABB08330 ID ABB08330 standard; protein; 67 AA.

XX ABB08330;

XX 18-JUN-2002 (first entry)

XX Bee venom protein Api m 6.01.

XX Bee venom: isoform; immunosuppressant; vaccine; Api m 6; immune response;
XX Bee venom hypersensitivity; antibody; protein purification; Api m 6.01;
XX immunotherapy; allergen.

XX Apis sp.

XX WO2001188085-A2.

XX 22-NOV-2001.

XX 16-FEB-2001; 2001WO-1801736.

XX 18-FEB-2000; 2000US-0506978.

XX (ECOL-) ECOLE POLYTECHNIQUE FEDERALE LAUSANNE.

XX Spertini F;

XX WPI; 2002-082988/11.

XX New bee venom polypeptides, useful for modulating immune responses e.g.
PT in individual hypersensitive to the venom and for identifying
PT individual at risk for bee venom hypersensitivity

```

XX PS Example 2; Page 26; 32pp; English.
XX CC
XX CC The present sequence is that of one the four isoforms of Api m 6,
XX CC designated Api m 6.01. The sequence represents the central amino acid
XX CC sequence shared by all four isoforms (see ABB08331, ABB08332 and
XX CC ABB08333). The specification describes a substantially pure polypeptide,
XX CC Api m 6, derived from bee venom and found in four isoforms. The proteins
XX CC of the invention have immunosuppressant activity and may form the basis
XX CC of a vaccine. Api m 6 is useful for modulating an immune response, i.e.
XX CC as an allergen for immunotherapy. The protein is useful for identifying
XX CC an individual at risk for bee venom hypersensitivity. The method
XX CC comprises administering Api m 6 to the individual and measuring an immune
XX CC response raised, where a detectable immune response indicates that the
XX CC individual is at risk for bee venom hypersensitivity. Antibodies specific
XX CC for Api m 6 are useful for purifying the protein.
XX SQ Sequence 67 AA;

Query Match 100.0%; Score 391; DB 23; Length 67;
Best Local Similarity 100.0%; Pred. No. 8e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGLGGRGKCPSEIFSRCDGRCQRCFPCNVVVKPLCIKICAPGCVCRGLYLRNKKKVC 60
Db 1 GFGGLGGRGKCPSEIFSRCDGRCQRCFPCNVVVKPLCIKICAPGCVCRGLYLRNKKKVC 60
QY 61 VPRSKCG 67
Db 61 VPRSKCG 67

RESULT 2
ABB08331
ID ABB08331 standard; protein: 69 AA.
XX AC
XX AC ABB08331;
XX DT 18-JUN-2002 (first entry)
XX DE Bee venom protein Api m 6.02.
XX KW Bee venom; isoform; immunosuppressant; vaccine; Api m 6; immune response;
XX KW bee venom hypersensitivity; antibody; protein purification; Api m 6.02;
XX KW immunotherapy; allergen.
XX OS
XX OS Apis sp.
XX PN W0200188085-A2.
XX PD 22-NOV-2001.
XX PF 16-FEB-2001; 2001WO-IB01736.
XX FT Misc-difference 68 /label= Pro, Leu
XX FT Misc-difference 69 /label= Pro, Leu
XX PN W0200188085-A2.
XX PD 22-NOV-2001.
XX PF 16-FEB-2001; 2001WO-IB01736.
XX PR 18-FEB-2000; 2000US-0506978.
XX PA (ECOL-) ECOLE POLYTECHNIQUE FEDERALE LAUSANNE.
XX PI Spertini F;
XX DR WPI; 2002-082988/11.
XX XX New bee venom polypeptides, useful for modulating immune responses e.g.
XX PT in individual hypersensitive to the venom and for identifying
XX PT individual at risk for bee venom hypersensitivity
XX PS Example 2; Page 26; 32pp; English.

XX PS Example 2; Page 26; 32pp; English.
XX CC
XX CC The present sequence is that of one the four isoforms of Api m 6,
XX CC designated Api m 6.02. All four isoforms share a common central amino
XX CC acid sequence shared by all four isoforms (see ABB08330, ABB08332 and
XX CC ABB08333). The specification describes a substantially pure polypeptide,
XX CC Api m 6, derived from bee venom and found in four isoforms. The proteins
XX CC of the invention have immunosuppressant activity and may form the basis
XX CC of a vaccine. Api m 6 is useful for modulating an immune response, i.e.
XX CC as an allergen for immunotherapy. The protein is useful for identifying
XX CC an individual at risk for bee venom hypersensitivity. The method
XX CC comprises administering Api m 6 to the individual and measuring an immune
XX CC response raised, where a detectable immune response indicates that the
XX CC individual is at risk for bee venom hypersensitivity. Antibodies specific
XX CC for Api m 6 are useful for purifying the protein.
XX SQ Sequence 69 AA;

Query Match 100.0%; Score 391; DB 23; Length 69;
Best Local Similarity 100.0%; Pred. No. 8.2e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGLGGRGKCPSEIFSRCDGRCQRCFPCNVVVKPLCIKICAPGCVCRGLYLRNKKKVC 60
Db 1 GFGGLGGRGKCPSEIFSRCDGRCQRCFPCNVVVKPLCIKICAPGCVCRGLYLRNKKKVC 60
QY 61 VPRSKCG 67
Db 61 VPRSKCG 67

RESULT 3
ABB08332
ID ABB08332 standard; protein: 71 AA.
XX AC
XX AC ABB08332;
XX DT 18-JUN-2002 (first entry)
XX DE Bee venom protein Api m 6.03.
XX KW Bee venom; isoform; immunosuppressant; vaccine; Api m 6; immune response;
XX KW bee venom hypersensitivity; antibody; protein purification; Api m 6.03;
XX KW immunotherapy; allergen.
XX OS
XX OS Apis sp.
XX PN W0200188085-A2.
XX PD 22-NOV-2001.
XX PF 16-FEB-2001; 2001WO-IB01736.
XX PR 18-FEB-2000; 2000US-0506978.
XX PA (ECOL-) ECOLE POLYTECHNIQUE FEDERALE LAUSANNE.
XX PI Spertini F;
XX WPI; 2002-082988/11.
XX PT New bee venom polypeptides, useful for modulating immune responses e.g.
XX PT in individual hypersensitive to the venom and for identifying
XX PT individual at risk for bee venom hypersensitivity
XX PS Example 2; Page 26; 32pp; English.

XX PS Example 2; Page 26; 32pp; English.
XX CC
XX CC The present sequence is that of one the four isoforms of Api m 6,
XX CC designated Api m 6.03. All four isoforms share a common central amino
XX CC acid sequence shared by all four isoforms (see ABB08330, ABB08331 and
XX CC ABB08333). The specification describes a substantially pure polypeptide,
XX CC Api m 6, derived from bee venom and found in four isoforms. The proteins
XX CC of the invention have immunosuppressant activity and may form the basis
XX CC of a vaccine. Api m 6 is useful for modulating an immune response, i.e.
XX CC as an allergen for immunotherapy. The protein is useful for identifying
XX CC an individual at risk for bee venom hypersensitivity. The method
XX CC comprises administering Api m 6 to the individual and measuring an immune
XX CC response raised, where a detectable immune response indicates that the
XX CC individual is at risk for bee venom hypersensitivity. Antibodies specific
XX CC for Api m 6 are useful for purifying the protein.
XX SQ Sequence 69 AA;

```

as an allergen for immunotherapy. The protein is useful for identifying an individual at risk for bee venom hypersensitivity. The method comprises administering Api m 6 to the individual and measuring an immune response raised, where a detectable immune response indicates that the individual is at risk for bee venom hypersensitivity. Antibodies specific for Api m 6 are useful for purifying the protein.

Sequence 71 AA;

Query Match 100.0%; Score 391; DB 23; Length 71;
Best Local Similarity 100.0%; Pred. No. 8.4e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGLGGGKCPNSIEFSRCDGRCQPCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60
|||||

Db 5 GGFGLGGGKCPNSIEFSRCDGRCQPCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 64
|||||

QY 61 VPRSKCG 67
|||||

Db 65 VPRSKCG 71
|||||

RESULT 4

ABB08333

ID ABB08333 standard; protein; 73 AA.

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

comprises administering Api m 6 to the individual and measuring an immune response raised, where a detectable immune response indicates that the individual is at risk for bee venom hypersensitivity. Antibodies specific for Api m 6 are useful for purifying the protein.

Sequence 73 AA;

Query Match 100.0%; Score 391; DB 23; Length 73;
Best Local Similarity 100.0%; Pred. No. 8.6e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGLGGGKCPNSIEFSRCDGRCQPCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60
|||||

Db 5 GGFGLGGGKCPNSIEFSRCDGRCQPCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 64
|||||

QY 61 VPRSKCG 67
|||||

Db 65 VPRSKCG 71
|||||

RESULT 5

AAV69209

ID AAY69209 standard; Protein; 92 AA.

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

comprises administering Api m 6 to the individual and measuring an immune response raised, where a detectable immune response indicates that the individual is at risk for bee venom hypersensitivity. Antibodies specific for Api m 6 are useful for purifying the protein.

Sequence 71 AA;

Query Match 100.0%; Score 391; DB 23; Length 71;
Best Local Similarity 100.0%; Pred. No. 8.4e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGLGGGKCPNSIEFSRCDGRCQPCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60
|||||

Db 5 GGFGLGGGKCPNSIEFSRCDGRCQPCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 64
|||||

QY 61 VPRSKCG 67
|||||

Db 65 VPRSKCG 71
|||||

RESULT 4

ABB08333

ID ABB08333 standard; protein; 73 AA.

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

XX ABB08333;

comprises administering Api m 6 to the individual and measuring an immune response raised, where a detectable immune response indicates that the individual is at risk for bee venom hypersensitivity. Antibodies specific for Api m 6 are useful for purifying the protein.

Sequence 73 AA;

Query Match 100.0%; Score 391; DB 23; Length 73;
Best Local Similarity 100.0%; Pred. No. 8.6e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGLGGGKCPNSIEFSRCDGRCQPCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60
|||||

Db 5 GGFGLGGGKCPNSIEFSRCDGRCQPCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 64
|||||

QY 61 VPRSKCG 67
|||||

Db 65 VPRSKCG 71
|||||

RESULT 5

AAV69209

ID AAY69209 standard; Protein; 92 AA.

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

XX AAY69209;

comprises administering Api m 6 to the individual and measuring an immune response raised, where a detectable immune response indicates that the individual is at risk for bee venom hypersensitivity. Antibodies specific for Api m 6 are useful for purifying the protein.

Sequence 71 AA;

Query Match 100.0%; Score 391; DB 23; Length 71;
Best Local Similarity 100.0%; Pred. No. 8.4e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGLGGGKCPNSIEFSRCDGRCQPCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60
|||||

Db 5 GGFGLGGGKCPNSIEFSRCDGRCQPCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 64
|||||

QY 61 V

responsible for genetic disorders or other traits and to assess biodiversity -

Claim 20; SEQ ID NO 52574; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at pub.int/pub/published.pct.sequences.

WPI; 1999-539569/45.

Screening an isolated protein for Nematode-extracted Anticoagulant Protein domains

Disclosure; Columns 142-144; 197pp; English.

The present sequence represents a nematode extracted anticoagulant protein (NAP). The protein has activity as an anticoagulant and/or serine protease inhibitor. The protein contains at least one NAP domain which has selective inhibitory activity for factor VIIa/TF. The specification describes a method for screening an isolated protein at least one domain for factor VIIa/TF selective inhibitory activity. The method comprises determining the time to clotting effected by a concentration of the isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo activated partial thromboplastin time (aPTT) assay; calculating prolongation of clotting effected by the isolated protein in each of the PT and aPTT assay, with respect to a baseline clotting value for each assay, where prolongation of clotting is calculated as fold elevation of clotting time relative to a baseline clotting value, where a doubling of clotting time is deemed a two-fold elevation; and calculating a PT to aPTT prolongation ratio, where a ratio at least one is indicative of factor VIIa/TF inhibitory activity. The method is useful for determining if a protein has factor VIIa/TF inhibitory activity.

Sequence 84 AA;

Query Match

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:05:12 : Search time 33 seconds
(without alignments)
76.587 Million cell updates/sec

Title: US-09-506-978-1
Perfect score: 391
Sequence: 1 GGFGGLGGKCPSENEFSR.....CRLGRLNKKKVCVPRSKG 67

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 174566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	106.5	27.2	5405	9 US-10-025-380-1116	Sequence 1116, Ap
2	106.5	27.2	5405	10 US-09-922-217-1116	Sequence 1116, Ap
3	98	25.1	2813	10 US-09-381-261A-1	Sequence 1, Appli
4	87.5	22.4	2813	10 US-09-886-900-2	Sequence 2, Appli
5	85	21.7	759	9 US-10-189-971-22	Sequence 22, Appli
6	85	21.7	1057	9 US-10-189-971-6	Sequence 6, Appli
7	85	21.7	1192	9 US-10-189-971-18	Sequence 18, Appli
8	85	21.7	1207	9 US-10-189-971-20	Sequence 20, Appli
9	85	21.7	1251	9 US-10-189-971-16	Sequence 16, Appli
10	85	21.7	1342	9 US-10-189-971-24	Sequence 24, Appli
11	85	21.7	1477	9 US-10-189-971-8	Sequence 8, Appli
12	85	21.7	1512	9 US-10-189-971-10	Sequence 10, Appli
13	85	21.7	1535	9 US-10-189-971-14	Sequence 14, Appli
14	85	21.7	1570	9 US-10-189-971-12	Sequence 12, Appli
15	85	21.7	1593	9 US-10-189-971-4	Sequence 4, Appli
16	85	21.7	1628	9 US-10-189-971-2	Sequence 2, Appli
17	83.5	21.4	355	10 US-09-826-212-14	Sequence 14, Appli
18	83.5	21.4	355	10 US-09-935-727-16	Sequence 16, Appli
19	82.5	21.1	469	10 US-09-925-301-1279	Sequence 1279, Ap

20	78.5	20.1	77	9 US-09-950-933A-46	Sequence 46, Appli
21	77.5	19.8	84	9 US-09-950-933A-54	Sequence 54, Appli
22	76.5	19.6	855	10 US-09-900-751-2	Sequence 2, Appli
23	75.5	19.3	349	10 US-09-826-212-13	Sequence 13, Appli
24	75.5	19.3	349	10 US-09-935-727-15	Sequence 15, Appli
25	75.5	19.3	5179	9 US-10-025-380-1068	Sequence 1068, Ap
26	75.5	19.3	5179	10 US-09-922-217-1068	Sequence 1068, Ap
27	75.5	19.3	5179	10 US-09-833-263-1068	Sequence 1068, Ap
28	72	18.4	366	10 US-09-205-658-103	Sequence 103, App
29	72	18.4	366	10 US-09-844-353A-103	Sequence 103, App
30	72	18.4	1367	9 US-09-870-759-120	Sequence 120, App
31	71	18.2	129	9 US-09-950-933A-43	Sequence 43, Appli
32	70	17.9	474	9 US-10-084-994-11	Sequence 11, Appli
33	68.5	17.5	3046	9 US-09-759-130B-441	Sequence 441, App
34	67	17.1	578	10 US-09-908-322-13	Sequence 13, Appli
35	66	16.9	1050	9 US-09-796-753-114	Sequence 114, App
36	66	16.9	1509	10 US-09-901-940-2	Sequence 2, Appli
37	65.5	16.8	96	9 US-09-950-933A-60	Sequence 60, Appli
38	65.5	16.8	689	10 US-09-071-838-2	Sequence 2, Appli
39	65	16.6	636	9 US-09-796-753-100	Sequence 100, App
40	65	16.6	636	9 US-09-796-753-124	Sequence 124, App
41	64	16.4	276	9 US-10-086-176A-5	Sequence 5, Appli
42	64	16.4	276	9 US-09-741-106-9	Sequence 9, Appli
43	64	16.4	276	10 US-09-766-778-1	Sequence 1, Appli
44	64	16.4	291	9 US-09-924-340-48	Sequence 48, Appli
45	64	16.4	291	9 US-09-924-340-52	Sequence 52, Appli

ALIGNMENTS

RESULT 1
US-10-025-380-1116
; Sequence 1116, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1116
; LENGTH: 5405
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-1116

Query Match 27.2%; Score 106.5; DB 9; Length 5405;
Best Local Similarity 36.8%; Pred. No. 0.027;
Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;
QY 11 KCPSENEFSRCDGRCQRCPCPNVVPKPLCKIKCAPCVCRLGRLNKKKVCVPRSKG 67
DB 1531 ECPNHNVELCADTCSLGSALSAPPQCDGACGCDGSGFLYN-GOACVPIQCCG 1586

Db 695 MDERGDCVPKAC 707

RESULT 4

US-09-886-900-2

Sequence 2, Application US/09886900

Patent No. US20020137051A1

GENERAL INFORMATION:

APPLICANT: Venta, Patrick J

Yuzbasiyan-Gurkan, Vilma

Schall, William D

Brewer, George J

TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND

FACTOR AND METHODS OF USE

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

STREET: 3445 Corporate Drive

CITY: Troy

STATE: Michigan

COUNTRY: USA

ZIP: 48098

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/886,900

FILING DATE: 21-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/896,449

FILING DATE: 18-JUL-1997

ATTORNEY/AGENT INFORMATION:

NAME: Smith, DeAnn F.

REFERENCE/DOCKET NUMBER: 2115-001226

TELECOMMUNICATION INFORMATION:

TELEPHONE: 248-641-1600

TELEFAX: 248-641-0270

TELEX: 287637

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2813 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-886-900-2

Query Match 22.4%; Score 87.5; DB 10; Length 2813;

Best Local Similarity 31.0%; Pred. No. 0.99;

Matches 18; Conservative 10; Mismatches 29; Indels 1; Gaps 1;

QY 9 RKGKPSNEIFSRCDGRCQRCPCNVVPRKPLCKIKICAPGCVRLGRLRNKKKVCVPRSKC 66

Db 292 RPACFAGMEYKECVSPCTRTQSLHVKVEVCQEQVDGSCPEGLLDEGH-CVGSAC 348

RESULT 5

US-10-189-971-22

Sequence 22, Application US/10189971

Publication No. US20030028907A1

GENERAL INFORMATION:

APPLICANT: Walke, D. Wade

APPLICANT: Scoville, John

APPLICANT: Turner, C. Alexander Jr.

TITLE OF INVENTION: NO. US20030028907A1el Human Kiellin-like Proteins and Polynucle

FILE OF INVENTION: Same

FILE REFERENCE: LEX-0360-USA

CURRENT APPLICATION NUMBER: US/10/189,971

CURRENT FILING DATE: 2002-07-03

PRIOR APPLICATION NUMBER: US 60/302,949


```
; Sequence 14, Application US/10189971
; Publication No. US20030028907A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030028907A1el Human Kiellin-like Proteins and Polynucleotides
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0360-USA
; CURRENT APPLICATION NUMBER: US/10/189,971
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/302,949
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/315,634
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1535
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-189-971-14
```

```
Query Match 21.7%; Score 85; DB 9; Length 1535;
Best Local Similarity 30.6%; Pred. No. 1;
Matches 19; Conservative 8; Mismatches 27; Indels 8; Gaps 4;

QY 12 CPSNE--IFSRCDGRCORCPNV-VP----KPLCIKICAPGCVCRGLYLRNKKKVCVPRS 64
Db 1450 CPLERGFVDECGPPCPTCFNQHIPLGELAAHCVRPCVPGCCOPAGLVEHAH-CIPPE 1508

QY 65 KC 66
Db 1509 AC 1510
```

```
RESULT 14
US-10-189-971-12
; Sequence 12, Application US/10189971
; Publication No. US20030028907A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030028907A1el Human Kiellin-like Proteins and Polynucleotides
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0360-USA
; CURRENT APPLICATION NUMBER: US/10/189,971
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/302,949
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/315,634
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1570
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-189-971-12
```

```
Query Match 21.7%; Score 85; DB 9; Length 1570;
Best Local Similarity 30.6%; Pred. No. 1;
Matches 19; Conservative 8; Mismatches 27; Indels 8; Gaps 4;

QY 12 CPSNE--IFSRCDGRCORCPNV-VP----KPLCIKICAPGCVCRGLYLRNKKKVCVPRS 64
Db 1485 CPLERGFVDECGPPCPTCFNQHIPLGELAAHCVRPCVPGCCOPAGLVEHAH-CIPPE 1543

QY 65 KC 66
Db 1544 AC 1545
```

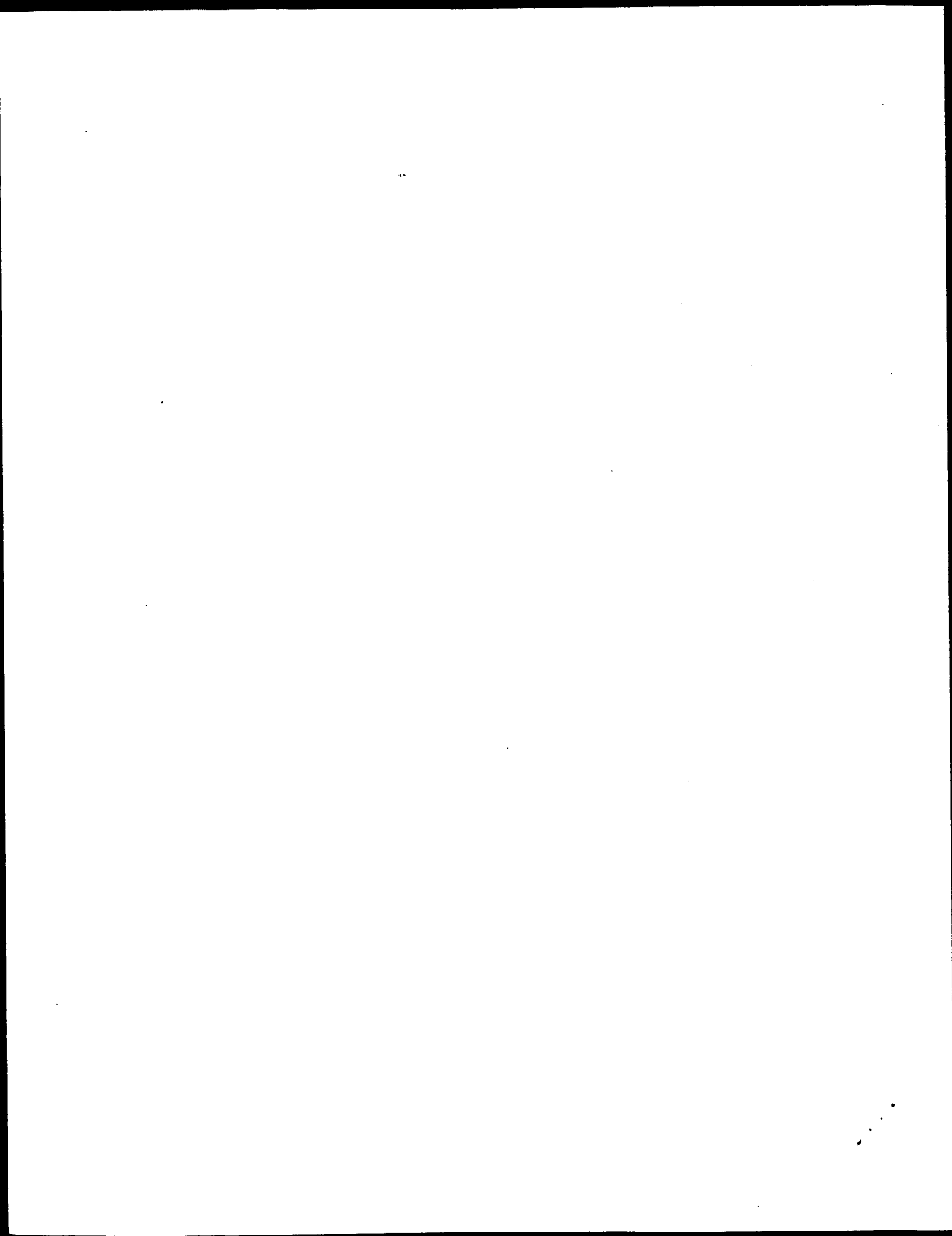
```
RESULT 15
US-10-189-971-4
; Sequence 4, Application US/10189971
; Publication No. US20030028907A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20030028907A1el Human Kiellin-like Proteins and Polynucleotides
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0360-USA
; CURRENT APPLICATION NUMBER: US/10/189,971
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/302,949
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/315,634
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1593
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-189-971-4
```

```
Query Match 21.7%; Score 85; DB 9; Length 1593;
Best Local Similarity 30.6%; Pred. No. 1.1;
Matches 19; Conservative 8; Mismatches 27; Indels 8; Gaps 4;

QY 12 CPSNE--IFSRCDGRCORCPNV-VP----KPLCIKICAPGCVCRGLYLRNKKKVCVPRS 64
Db 1508 CPLERGFVDECGPPCPTCFNQHIPLGELAAHCVRPCVPGCCOPAGLVEHAH-CIPPE 1566

QY 65 KC 66
Db 1567 AC 1568
```

Search completed: February 26, 2003, 15:10:00
Job time : 36 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:04:07 ; Search time 141 seconds
(without alignments)
306.363 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGFGLGGRGKCPSEIFSR.....CRLGYLRNKKVCVPRSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	391	100.0	67	19	US-09-506-978-1
2	391	100.0	67	25	US-10-174-151-1
3	391	100.0	67	26	US-10-204-145-1
4	391	100.0	69	19	US-09-506-978-2
5	391	100.0	69	25	US-10-174-151-2
6	391	100.0	69	26	US-10-204-145-2

7	391	100.0	71	19	US-09-506-978-3	Sequence 3, Appli
8	391	100.0	71	25	US-10-174-151-3	Sequence 3, Appli
9	391	100.0	71	26	US-10-204-145-3	Sequence 3, Appli
10	391	100.0	73	19	US-09-506-978-4	Sequence 4, Appli
11	391	100.0	73	25	US-10-174-151-4	Sequence 4, Appli
12	391	100.0	73	26	US-10-204-145-4	Sequence 4, Appli
13	391	100.0	92	1	PCT-US99-21077-2	Sequence 2, Appli
14	113.5	29.0	62	27	US-60-142-896-1491	Sequence 1491, Ap
15	113.5	29.0	62	27	US-60-145-138-701	Sequence 701, App
16	113.5	29.0	197	16	US-09-270-767-43953	Sequence 43953, A
17	111.5	28.5	166	27	US-60-360-039-6314	Sequence 6314, Ap
18	109	27.9	1537	27	US-60-230-445-1775	Sequence 1775, Ap
19	109	27.9	2554	27	US-60-230-445-1881	Sequence 1881, Ap
20	109	27.9	2593	27	US-60-207-583-584	Sequence 584, App
21	109	27.9	2601	24	US-10-016-248-59	Sequence 59, Appl
22	109	27.9	2659	27	US-60-207-583-632	Sequence 632, App
23	108	27.6	164	27	US-60-138-684-918	Sequence 918, App
24	108	27.6	1010	27	US-60-147-189-857	Sequence 857, App
25	108	27.6	3644	27	US-60-167-324-871	Sequence 871, App
26	108	27.6	3644	27	US-60-171-625-561	Sequence 561, App
27	108	27.6	3644	27	US-60-173-386-841	Sequence 841, App
28	108	27.6	3644	27	US-60-175-871-955	Sequence 955, App
29	108	27.6	3644	27	US-60-184-775-869	Sequence 869, App
30	108	27.6	3644	27	US-60-191-637-41010	Sequence 41010, A
31	108	27.6	3644	27	US-60-191-700-930	Sequence 930, App
32	108	27.6	3843	20	US-09-614-150-41379	Sequence 41379, A
33	106.5	27.2	76	27	US-60-192-739-2374	Sequence 2374, Ap
34	106.5	27.2	2957	1	PCT-US01-08631-52573	Sequence 52573, A
35	106.5	27.2	3014	27	US-60-258-273-160	Sequence 160, App
36	106.5	27.2	5405	23	US-09-922-217-1116	Sequence 1116, Ap
37	106.5	27.2	5405	24	US-10-025-380-1116	Sequence 1116, Ap
38	101.5	26.0	77	27	US-60-196-174-1252	Sequence 1252, Ap
39	101.5	26.0	91	27	US-60-162-243-571	Sequence 571, App
40	101.5	26.0	91	27	US-60-162-247-3514	Sequence 3514, Ap
41	101.5	26.0	126	27	US-60-177-571-4507	Sequence 4507, Ap
42	101.5	26.0	126	27	US-60-177-646-3718	Sequence 3718, Ap
43	101.5	26.0	7337	1	PCT-US01-08631-52575	Sequence 52575, A
44	100.5	25.7	135	27	US-60-196-174-1284	Sequence 1284, Ap
45	100.5	25.7	735	1	PCT-US01-08631-52574	Sequence 52574, A

ALIGNMENTS

RESULT 1

US-09-506-978-1

; Sequence 1, Application US/09506978

; GENERAL INFORMATION:

; APPLICANT: Spertini, Francois

; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF

; FILE REFERENCE: 18519-001

; CURRENT APPLICATION NUMBER: US/09/506,978

; CURRENT FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 67

; TYPE: PRT

; ORGANISM: Apis mellifera

US-09-506-978-1

Query Match

Best Local Similarity 100.0%; Score 391; DB 19; Length 67;

Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGLGGRGKCPSEIFSRCDRCORFCPNVVPKPLCIKICAPGCVCLGYLRNKKVC 60

Db 1 GGFGLGGRGKCPSEIFSRCDRCORFCPNVVPKPLCIKICAPGCVCLGYLRNKKVC 60

QY 61 VPRSKCG 67

Db 61 VPRSKCG 67

```
RESULT 2
US-10-174-151-1
; Sequence 1, Application US/10174151
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/10/174,151
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Apis mellifera
US-10-174-151-1
Query Match      100.0%; Score 391; DB 25; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.1e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60
Db 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60

Qy 61 VPRSKCG 67
Db 61 VPRSKCG 67

RESULT 3
US-10-204-145-1
; Sequence 1, Application US/10204145
; GENERAL INFORMATION:
; APPLICANT: Ecole Polytechnique Federale de Lausanne
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001-064 20349-543
; CURRENT APPLICATION NUMBER: US/10/204,145
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Apis mellifera
US-10-204-145-1
Query Match      100.0%; Score 391; DB 26; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.1e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60
Db 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60

Qy 61 VPRSKCG 67
Db 61 VPRSKCG 67

RESULT 4
US-09-506-978-2
; Sequence 2, Application US/09506978
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
```

```
; CURRENT APPLICATION NUMBER: US/09/506,978
; CURRENT FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Apis mellifera
; FEATURE:
; OTHER INFORMATION: Where XX at position 68 and 69 is either PL or LP.
US-09-506-978-2
Query Match      100.0%; Score 391; DB 19; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.2e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60
Db 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60

Qy 61 VPRSKCG 67
Db 61 VPRSKCG 67

RESULT 5
US-10-174-151-2
; Sequence 2, Application US/10174151
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/10/174,151
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Apis mellifera
; FEATURE:
; OTHER INFORMATION: Where XX at position 68 and 69 is either PL or LP.
US-10-174-151-2
Query Match      100.0%; Score 391; DB 25; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.2e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60
Db 1 GFGGLGGRGKCPSPNEIFSRCDGRQFCPNVVPKPLCIKICAPGCVCRGLYLRNKKKVC 60

Qy 61 VPRSKCG 67
Db 61 VPRSKCG 67

RESULT 6
US-10-204-145-2
; Sequence 2, Application US/10204145
; GENERAL INFORMATION:
; APPLICANT: Ecole Polytechnique Federale de Lausanne
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001-064 20349-543
; CURRENT APPLICATION NUMBER: US/10/204,145
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
```

```

; SEQ ID NO 2
; LENGTH: 69
; TYPE: PRT
; ORGANISM:
; FEATURE:
; OTHER INFO:
US-10-204-145-

```

Query Match 100.0%; Score 391; DB 26; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.2e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 VPRSKCG 67
| | | | |
Db 61 VPRSKCG 67

```

RESULT 7
US-09-506-978-3
; Sequence 3, Application US/09506978
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/09/506, 978
; CURRENT FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Apis mellifera
US-09-506-978-3

```

```

RESULT 8
US-10-174-151-3
; Sequence 3, Application US/10174151
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/10/174,151
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Apis mellifera
US-10-174-151-3

```

QY	1	GGFGGLGGRGKCP	SNEI	ESRCDGRCQRF	CPNV	VPKPLCTKICAPG	VCVLGRNKKVC	60
Db	5	GGFGGLGGRGKCP	SNEI	ESRCDGRCQRF	CPNV	VPKPLCTKICAPG	VCVLGRNKKVC	64
QY	61	VPRSKG	67					
Db	65	VPRSKG	71					

```

RESULT 9
US-10-204-145-3
; Sequence 3, Application US/10204145
; GENERAL INFORMATION:
; APPLICANT: Ecole Polytechnique Federale de Lausanne
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001-064 20349-543
; CURRENT APPLICATION NUMBER: US/10/204,145
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Apis mellifera
US-10-204-145-3

```

```

RESULT 10
US-09-506-978-4
; Sequence 4, Application US/09506978
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/09/506,978
; CURRENT FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin ver. 2.0
; SEQ ID NO 4
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Apis mellifera
; FEATURE:
; OTHER INFORMATION: Where XX at position 72 and 73 is either PL or LP.
US-09-506-978-4

```

Query Match	100.0%;	Score 391;	DB 19;	Length 73;
Best Local Similarity	100.0%;	Pred. No. 3 4e-32;		
Matches	67;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	GGFGGIGRGKCP	SNEIFSRCDGRCQRF	CPNVVPKLCIKICAPGVCRLGYLRNKKVC 60
Db	5	GGFGGIGRGKCP	SNEIFSRCDGRCQRF	CPNVVPKLCIKICAPGVCRLGYLRNKKVC 64
QY	61	VPRSKCG	67	
Db	65	VPRSKCG	71	

```

; Sequence 2, Application PC/TUS9921077
; GENERAL INFORMATION:
; APPLICANT: Cui, Xiangmin
; APPLICANT: Lu, Yufeng
; APPLICANT: Pan Pacific Pharmaceutical, Inc.
; TITLE OF INVENTION: Useful Properties of a Bee Venom Protein and Gene
; TITLE OF INVENTION: Encoding Same
; FILE REFERENCE: 019049-000200PC
; CURRENT APPLICATION NUMBER: PCT/US99/21077
; CURRENT FILING DATE: 1999-09-13
; EARLIER APPLICATION NUMBER: US 60/100,172
; EARLIER FILING DATE: 1998-09-14
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Apis mellifera
; PCT-US99-21077-2

Query Match      100.0%; Score 391; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 4.1e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGGLGGGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60
    |||||
DB 26 GFGGGLGGGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 85

QY 61 VPRSKCG 67
    |||||
DB 86 VPRSKCG 92

RESULT 14
US-60-142-896-1491
; Sequence 1491, Application US/60142896
; GENERAL INFORMATION:
; APPLICANT: Kerlavage, Anthony
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING SECRETED PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000052
; CURRENT APPLICATION NUMBER: US/60/142,896
; CURRENT FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 1584
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1491
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Drosophila
; US-60-142-896-1491

Query Match      29.0%; Score 113.5; DB 27; Length 62;
Best Local Similarity 36.2%; Pred. No. 0.0008;
Matches 21; Conservative 8; Mismatches 26; Indels 3; Gaps 2;

QY 11 KCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGYLRNKKKVCVPRSKC 66
    |||||
DB 4 RCPANETFLAGGPDQTECA-TLCKPCLVHRIRCPDGCYCNCNKGFAARNAAGTCIPLRRC 60

RESULT 15
US-60-145-138-701
; Sequence 701, Application US/60145138
; GENERAL INFORMATION:
; APPLICANT: Bonazza, Vivien
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING SECRETED PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000062
; CURRENT APPLICATION NUMBER: US/60/145,138
; CURRENT FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 701

; Sequence 4, Application US/10174151
; GENERAL INFORMATION:
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001
; CURRENT APPLICATION NUMBER: US/10/174,151
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US/09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Apis mellifera
; FEATURE:
; OTHER INFORMATION: Where XX at position 72 and 73 is either PL or LP.
; US-10-174-151-4

Query Match      100.0%; Score 391; DB 25; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.4e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGGLGGGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60
    |||||
DB 5 GFGGGLGGGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 64

QY 61 VPRSKCG 67
    |||||
DB 65 VPRSKCG 71

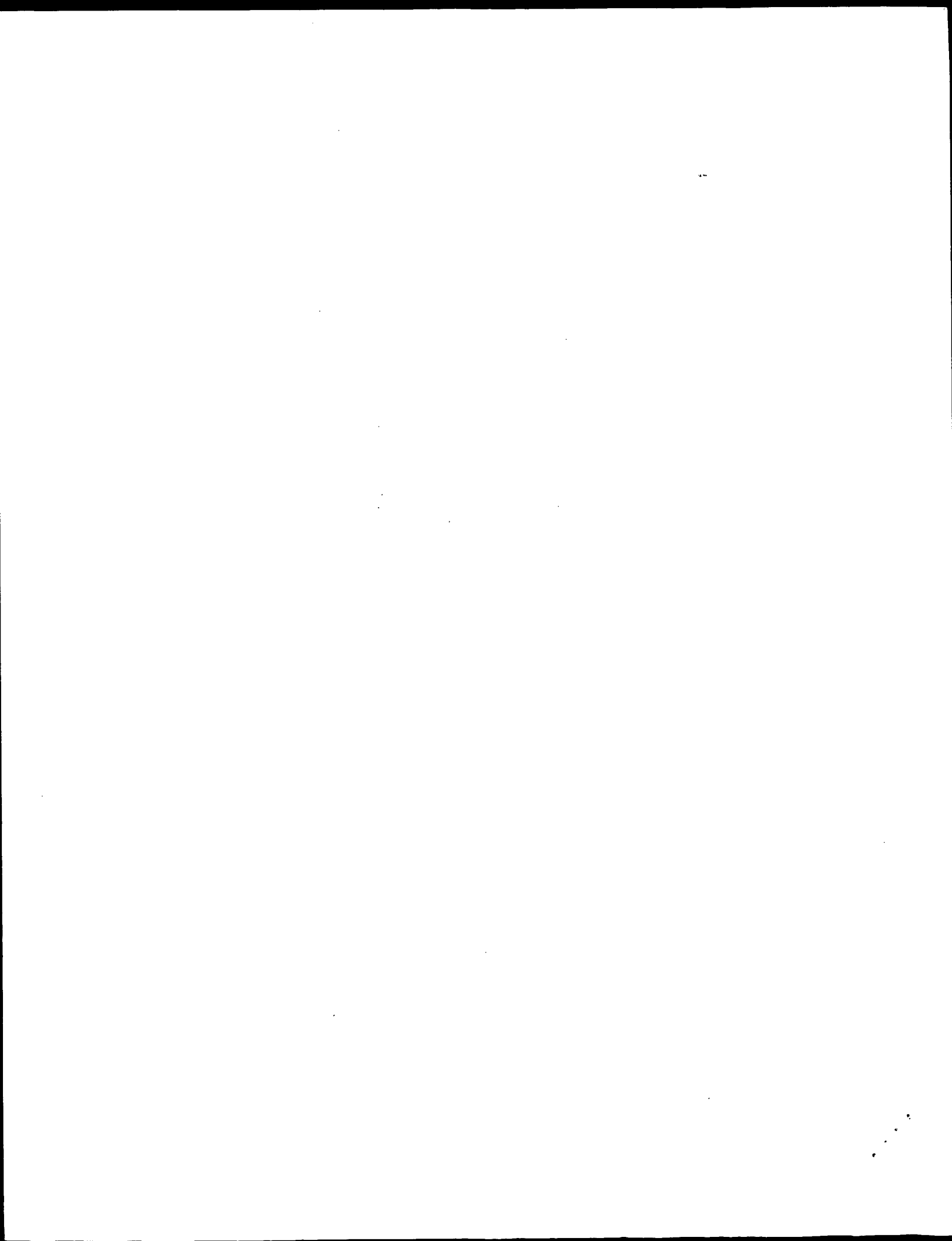
RESULT 12
US-10-204-145-4
; Sequence 4, Application US/10204145
; GENERAL INFORMATION:
; APPLICANT: Ecole Polytechnique Federale de Lausanne
; APPLICANT: Spertini, Francois
; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 18519-001-064 20349-543
; CURRENT APPLICATION NUMBER: US/10/204,145
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/506,978
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Apis mellifera
; FEATURE:
; OTHER INFORMATION: Where XX at position 72 and 73 is either PL or LP.
; US-10-204-145-4

Query Match      100.0%; Score 391; DB 26; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.4e-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GFGGGLGGGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 60
    |||||
DB 5 GFGGGLGGGKCPSPNEIFSRCDGRCQRCPCPNVVPKPLCIKICAPGCVRLGYLRNKKKVC 64

QY 61 VPRSKCG 67
    |||||
DB 65 VPRSKCG 71

RESULT 13
PCT-US99-21077-2
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:04:32 ; Search time 23 seconds
(without alignments)
266.242 Million cell updates/sec

Title: US-09-506-978-1
Perfect score: 391
Sequence: 1 GFGGLGGRGKCPNSIEFSR.....CRLGYLRNKKKVCVPSKCG 67

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 486122

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New :

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match %	Length	DB ID	Description
1	119	30.4	67	6	US-10-038-854-185	Sequence 185, App
2	113	28.9	5374	6	US-10-028-248A-75	Sequence 75, App
3	113	28.9	5374	6	US-10-107-782-75	Sequence 75, App
4	113	28.9	5376	6	US-10-028-248A-74	Sequence 74, App
5	113	28.9	5376	6	US-10-107-782-74	Sequence 74, App
6	109	27.9	2601	6	US-10-028-248A-76	Sequence 76, App
7	109	27.9	2601	6	US-10-107-782-76	Sequence 76, App
8	106.5	27.2	4360	5	US-09-724-676-66993	Sequence 66993, A
9	106.5	27.2	4360	5	US-09-724-676A-66993	Sequence 66993, A
10	98	25.1	2813	1	PCT-US02-04915-149	Sequence 149, App
11	97.5	24.9	111	1	PCT-US02-40655-379	Sequence 379, App
12	90	23.0	4219	6	US-10-085-198-2	Sequence 2, App
13	81	20.7	1082	6	US-10-264-237-2725	Sequence 2725, App
14	81	20.7	1458	1	PCT-US02-13209-25	Sequence 25, App
15	81	20.7	2212	6	US-10-028-248A-43	Sequence 24, App
16	81	20.7	2212	6	US-10-107-782-43	Sequence 43, App
17	81	20.7	2212	6	US-10-028-248A-42	Sequence 42, App
18	81	20.7	2570	6	US-10-107-782-42	Sequence 42, App
19	81	20.7	2570	6	US-10-263-929-117	Sequence 117, App
20	78	19.9	1568	6	US-10-072-012-355	Sequence 355, App
21	76.5	19.6	855	6	US-10-072-012-413	Sequence 413, App
22	76.5	19.6	855	6	US-10-072-012-413	Sequence 413, App
23	76.5	19.6	902	6	US-10-333-743-3	Sequence 3, App
24	75.5	19.3	5179	7	US-60-438-735-151	Sequence 151, App
25	73.5	18.8	855	6	US-10-072-012-356	Sequence 356, App
26	73.5	18.8	855	6	US-10-072-012-414	Sequence 414, App

27	73.5	18.8	855	6	US-10-072-012-417	Sequence 417, App
28	73	18.7	319	6	US-10-017-161-1346	Sequence 1346, App
29	73	18.7	969	6	US-10-052-648A-34	Sequence 34, App
30	73	18.7	969	6	US-10-052-648A-35	Sequence 35, App
31	73	18.7	969	6	US-10-055-877-214	Sequence 214, App
32	72	18.4	366	5	US-09-963-693-103	Sequence 103, App
33	72	18.4	1367	5	US-09-751-708A-120	Sequence 120, App
34	72	18.4	1367	5	US-09-949-001-18	Sequence 18, App
35	72	18.4	1376	6	US-10-276-774-2170	Sequence 2170, App
36	72	18.4	1377	5	US-09-949-001-21	Sequence 21, App
37	72	18.4	1664	6	US-10-055-877-212	Sequence 212, App
38	71	18.2	1511	6	US-10-228-206-11	Sequence 11, App
39	71	18.2	1581	5	US-09-949-002-414	Sequence 414, App
40	71	18.2	1587	5	US-09-949-002-354	Sequence 354, App
41	71	18.2	1620	6	US-10-055-877-213	Sequence 213, App
42	70	17.9	1393	6	US-10-312-352-21	Sequence 21, App
43	70	17.9	1398	6	US-10-055-877-46	Sequence 46, App
44	70	17.9	1403	6	US-10-055-877-52	Sequence 52, App
45	70	17.9	1404	6	US-10-055-877-44	Sequence 44, App

ALIGNMENTS

RESULT 1

US-10-038-854-185

; Sequence 185, Application US/10038854

; GENERAL INFORMATION:

; APPLICANT: Spytek, Kimberly A

; APPLICANT: Li, Li

; APPLICANT: Wolenc, Adam R

; APPLICANT: Vernet, Corine

; APPLICANT: Eisen, Andrew J

; APPLICANT: Liu, Xiaohong

; APPLICANT: Malyankar, Uriel M

; APPLICANT: Shimkets, Richard A

; APPLICANT: Tchernev, Velizar

; APPLICANT: Spaderna, Steven K

; APPLICANT: Gorman, Linda

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Patturajan, Meera

; APPLICANT: Gusev, Vladimir Y

; APPLICANT: Gangolli, Esha A

; APPLICANT: Guo, Xiaojia S

; APPLICANT: Shenoy, Suresh G

; APPLICANT: Rastelli, Luca J

; APPLICANT: Casman, Stacie J

; APPLICANT: Boldog, Ferenc

; APPLICANT: Burgess, Catherine E

; APPLICANT: Edinger, Shlomit R

; APPLICANT: Ellerman, Karen

; APPLICANT: Gunther, Erik

; APPLICANT: Smithson, Glennda

; APPLICANT: Millet, Isabelle

; APPLICANT: MacDougall, John R

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-230

; CURRENT APPLICATION NUMBER: US/10/038,854

; CURRENT FILING DATE: 2003-01-22

; PRIOR APPLICATION NUMBER: 60/258,928

; PRIOR FILING DATE: 2000-12-29

; PRIOR APPLICATION NUMBER: 60/259,415

; PRIOR FILING DATE: 2001-01-02

; PRIOR APPLICATION NUMBER: 60/259,785

; PRIOR FILING DATE: 2001-01-04

; PRIOR APPLICATION NUMBER: 60/269,814

; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/279,832

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,833

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,863

; PRIOR FILING DATE: 2001-03-29

```
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/285,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 185
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-854-185

Query Match      30.4%; Score 119; DB 6; Length 67;
Best Local Similarity 37.3%; Pred. No. 1.3e-05;
Matches 25; Conservative 7; Mismatches 23; Indels 12; Gaps 3;

QY 12 CPSNEIFSRCDGRCQRCFCPN-----VVP-----KPLCIKICAPGCVGRLGRLNKK-KV 59
Db 1 CPANQYTECGSPSCPSNPDGPLETTPCEGTSKVPSTCKECCVCGVVRNNDGDK 60
QY .60 CVPRSKC 66
Db 61 CVPRSEC 67

RESULT 2
US-10-028-248A-75
; Sequence 75, Application US/10028248A
; GENERAL INFORMATION:
; APPLICANT: Shimkets, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides and Methods of Use
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
```

```
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 5374
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-028-248A-75

Query Match      28.9%; Score 113; DB 6; Length 5374;
Best Local Similarity 37.5%; Pred. No. 0.0014;
Matches 24; Conservative 8; Mismatches 22; Indels 10; Gaps 3;

QY 11 KPSPNEIFSRCDGRCQRCFCPNVVPKPLC-----IKICAPGCVGRLGRLNKKKVCVPR 63
Db 4741 KCPANSLYTHCLPTCLPSCSN--PDGRCEGTSHKAPSTCREGCVQPGYLLN-KDTCVHK 4797
QY 64 SKCG 67
Db 4798 NQCG 4801

RESULT 3
US-10-107-782-75
; Sequence 75, Application US/10107782
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Millet, Charles,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREO
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 75
; LENGTH: 5374
```


; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66993
; LENGTH: 4360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66993

Query Match 27.2%; Score 106.5; DB 5; Length 4360;
Best Local Similarity 36.8%; Pred. No. 0.0057;
Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;
QY 11 KCPSNEIFSRCDGRCQRCFPCNVVPKPLCIKICAPGCVCRGLYLRNKKKVCVPRSKCG 67
Db 1531 ECPNPNHYELCADTCSLGCSALSAPPQCDGCAEGCQCDGFLYN-GQACVPIQQCG 1586

RESULT 9
US-09-724-676A-66993
; Sequence 66993, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66993
; LENGTH: 4360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-66993

Query Match 27.2%; Score 106.5; DB 5; Length 4360;
Best Local Similarity 36.8%; Pred. No. 0.0057;
Matches 21; Conservative 6; Mismatches 29; Indels 1; Gaps 1;
QY 11 KCPSNEIFSRCDGRCQRCFPCNVVPKPLCIKICAPGCVCRGLYLRNKKKVCVPRSKCG 67
Db 1531 ECPNPNHYELCADTCSLGCSALSAPPQCDGCAEGCQCDGFLYN-GQACVPIQQCG 1586

RESULT 10
PCT-US02-04915-149
; Sequence 149, Application PC/TUS0204915
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynnne, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; FILE REFERENCE: 018501-006200PC
; CURRENT APPLICATION NUMBER: PCT/US02/04915
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/285,475
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29

; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 149
; LENGTH: 2813
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-04915-149

Query Match 25.1%; Score 98; DB 1; Length 2813;
Best Local Similarity 26.0%; Pred. No. 0.033;
Matches 19; Conservative 13; Mismatches 27; Indels 14; Gaps 2;

QY 8 GRG-----KCPNNEIFSRCDGRCQRCFPCNV-VKPLCIKICAPGCVCRGLYL 53
Db 635 GRGVRVANREPGRCEINCPKGGQVILQCGTFCNLTCRSLSYPDCEACLEGCFCPPGLY 694
QY 54 RNKKKVCVPRSKC 66
Db 695 MDERGDCVPKAC 707

RESULT 11
PCT-US02-40655-379
; Sequence 379, Application PC/TUS0240655
; GENERAL INFORMATION:
; APPLICANT: diaDexus, Inc.
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO PANCREATIC SPECIFIC GENES
; FILE REFERENCE: DEX-0376
; CURRENT APPLICATION NUMBER: PCT/US02/40655
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/342,768
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 543
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 379
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapien
PCT-US02-40655-379

Query Match 24.9%; Score 97.5; DB 1; Length 111;
Best Local Similarity 28.6%; Pred. No. 0.0035;
Matches 16; Conservative 13; Mismatches 26; Indels 1; Gaps 1;

QY 12 CPSNEIFSRCDGRCQRCFPCNV-VKPLCIKICAPGCVCRGLYLRNKKKVCVPRSKC 66
Db 34 CPKGQVYLQCGTFCNLTCRSLSYPDCEACLEGCFCPPGLYMDERGDCVPKAC 89

RESULT 12
US-10-085-198-2
; Sequence 2, Application US/10085198
; GENERAL INFORMATION:
; APPLICANT: Alsobrook et al.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-279
; CURRENT APPLICATION NUMBER: US/10/085,198
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/271,646
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 60/276,401
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,981
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/312,858
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/271,840
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/277,324
; PRIOR FILING DATE: 2001-03-20

```

RESULT 14
PCT-US02-13209-25
: Sequence 25, Application PC/TUS0213209
: GENERAL INFORMATION:
: APPLICANT: WEIGEL,
: APPLICANT: WEIGEL, JANET A.
: TITLE OF INVENTION: METHODS OF USING A HYDROLYZABLE POLYMER
: FILE REFERENCE: 619758-B/JP/286, 468
: CURRENT APPLICATION NUMBER: PCT/US02/13209

```

```

Query Match      20.7%; Score 81; DB 1; Length 1460;
Best Local Similarity 31.9%; pred. No. 1.3;
Matches 23; Conservative 9; Mismatches 22; Indels 18; Gaps 4;

Qy 5 GLGGRGKPSNEIFS--RCDGRQRFQNNVVKPLCIKICAPGCVCR-----LGYL R 54
      ||||| | : || : : || : ||| ||| |||
Db 922 GLGSGSCFCDGWTGPRCEVQLEL-----QPVTCPPEAVCAVCRAGNSEC SLGY-E 973

Qy 55 NKKKVCVPRSKC 66
      : ||| |
Db 974 GDGRVCTVADLC 985

```

Search completed: February 26, 2003, 15:09:19
Job time : 25 secs

GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,380
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/268
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-465-380-59
Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;
QY 11 KCPSEIFSRCDGR-CQRFQ-----PNVVKPLCIKICAPGCVCRGLYLNKK 57
DB 5 QCGENEKYDSCGSKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFYRNKD 61
QY 58 KVCVPRSKC 66
DB 62 DKCVSAEDC 70
RESULT 6
US-08-486-397-59
Sequence 59, Application US/08486397
Patent No. 5866542
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/268
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-59
Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;
QY 11 KCPSEIFSRCDGR-CQRFQ-----PNVVKPLCIKICAPGCVCRGLYLNKK 57
DB 5 QCGENEKYDSCGSKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFYRNKD 61
QY 58 KVCVPRSKC 66
DB 62 DKCVSAEDC 70
RESULT 7
US-08-486-399-59
Sequence 59, Application US/08486399
Patent No. 5866543
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.

NUMBER OF SEQUENCES: 357
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,397
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/268
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-486-397-59
Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;
QY 11 KCPSEIFSRCDGR-CQRFQ-----PNVVKPLCIKICAPGCVCRGLYLNKK 57
DB 5 QCGENEKYDSCGSKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFYRNKD 61
QY 58 KVCVPRSKC 66
DB 62 DKCVSAEDC 70
RESULT 7
US-08-486-399-59
Sequence 59, Application US/08486399
Patent No. 5866543
GENERAL INFORMATION:
APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jespers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.

```

;
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,399
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/270
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-486-399-59

Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPSEIFSRCDGR-CQRF-----PNVVPKPLCIKICAPGCVRLGVLNRKK 57
Db 5 QCGENEKYDCSGSKCKKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFRNKD 61

QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

RESULT 8
US-08-461-965-59
; Sequence 59, Application US/08461965
; Patent No. 5872098
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/461,965
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 210/243
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-461-965-59

Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPSEIFSRCDGR-CQRF-----PNVVPKPLCIKICAPGCVRLGVLNRKK 57
Db 5 QCGENEKYDCSGSKCKKCKYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFRNKD 61

QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

RESULT 9
US-08-634-641-59
; Sequence 59, Application US/08634641
; Patent No. 5955294
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George P. Vlasuk
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Mensens, Joris Hilda Llieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/634,641
; FILING DATE: April 19, 1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13231

```

```
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 219/136
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-08-634-641-59

Query Match 25.6%; Score 100; DB 2; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KPSNEIFSRCDGR-CORFC-----PNVVPKPLCIKICAPGCVCRGLGYNKK 57
Db 5 QCGENEKYDCSGSKEDCKKCYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFGYRNKD 61

QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

RESULT 10
US-09-249-471-59
; Sequence 59, Application US/09249471
; Patent No. 6040441
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Llieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: California
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
```

```
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,471
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-09-249-471-59

Query Match 25.6%; Score 100; DB 3; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KPSNEIFSRCDGR-CORFC-----PNVVPKPLCIKICAPGCVCRGLGYNKK 57
Db 5 QCGENEKYDCSGSKEDCKKCYDGVVEEDDEPNV---PCLVRVCHQDCVCEGFGYRNKD 61

QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

RESULT 11
US-09-249-472-59
; Sequence 59, Application US/09249472
; Patent No. 6046318
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Llieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Gansemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
; TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
```

```
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,472
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-09-249-472-59

Query Match 25.6%; Score 100; DB 3; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

Qy 11 KCPNFIKRCDCGR-CQRF-----PNNVVKPLCIKICAPGCVCRGLYLRNKK 57
:| || : | : | : | || | :|| | :||
Db 5 QCGENEKYDCSGSKCKKCKYDGVBEEDDEPNV---PCLVRVCHQDCVCEGFEYRNKD 61

Qy 58 KVCVPRSKC 66
|| |
Db 62 DKCVSAEDC 70

RESULT 12
US-09-249-451-59
; Sequence 59, Application US/09249451
; Patent No. 6087487
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
; APPLICANT: Messens, Joris Hilda Lieven
; APPLICANT: Lauwereys, Marc Josef
; APPLICANT: Laroche, Yves Rene
; APPLICANT: Jespers, Laurent Stephane
; APPLICANT: Ganssemans, Yannick Georges Jozef
; APPLICANT: Moyle, Matthew
; APPLICANT: Bergum, Peter W.
; TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE

; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/249,472
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/809,455
; FILING DATE: April 17, 1997
; APPLICATION NUMBER: PCT/US95/13231
; FILING DATE: October 17, 1995
; APPLICATION NUMBER: 08/486,399
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/486,397
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/465,380
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/461,965
; FILING DATE: June 5, 1995
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 216/270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Ancylostoma caninum
; US-09-249-472-59

Query Match 25.6%; Score 100; DB 3; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

Qy 11 KCPNFIKRCDCGR-CQRF-----PNNVVKPLCIKICAPGCVCRGLYLRNKK 57
:| || : | : | : | || | :|| | :||
Db 5 QCGENEKYDCSGSKCKKCKYDGVBEEDDEPNV---PCLVRVCHQDCVCEGFEYRNKD 61

Qy 58 KVCVPRSKC 66
|| |
Db 62 DKCVSAEDC 70

RESULT 13
US-08-809-455-59
; Sequence 59, Application US/08809455
; Patent No. 6090916
; GENERAL INFORMATION:
; APPLICANT: Vlasuk, George Phillip
; APPLICANT: Stanssens, Patrick Eric Hugo
```

APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Gansemans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
APPLICANT: Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,455
FILING DATE: April 17, 1997
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-08-809-455-59

Query Match 25.6%; Score 100; DB 3; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPNSNIFSRCDGR-CQRF-----PNVVPKPLCIKICAPGCVRLGLYLRNKK 57
Db 5 QCGENKDYSCGSKCKDKKYDGVVEEDDEPNV---PCLVRVCHQDCVCEEGFYRNKD 61

QY 58 KVCVPRSKC 66

Db 62 DKCVSAEDC 70

RESULT 14

US-09-249-461-59
Sequence 59, Application US/09249461
Patent No. 6096877
GENERAL INFORMATION:
APPLICANT: Vlasuk, George Phillip
APPLICANT: Staussens, Patrick Eric Hugo
APPLICANT: Messens, Joris Hilda Lieven
APPLICANT: Lauwereys, Marc Josef
APPLICANT: Laroche, Yves Rene
APPLICANT: Jespers, Laurent Stephane
APPLICANT: Gansemans, Yannick Georges Jozef
APPLICANT: Moyle, Matthew
APPLICANT: Bergum, Peter W.
TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
NUMBER OF SEQUENCES: 356
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/249,461
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/809,455
FILING DATE: April 17, 1997
APPLICATION NUMBER: PCT/US95/13231
FILING DATE: October 17, 1995
APPLICATION NUMBER: 08/486,399
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/486,397
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/465,380
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/461,965
FILING DATE: June 5, 1995
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 216/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Ancylostoma caninum
US-09-249-461-59

Query Match 25.6%; Score 100; DB 3; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPNSNIFSRCDGR-CQRF-----PNVVPKPLCIKICAPGCVRLGLYLRNKK 57
Db 5 QCGENKDYSCGSKCKDKKYDGVVEEDDEPNV---PCLVRVCHQDCVCEEGFYRNKD 61

Wed Feb 26 15:12:38 2003

US-09-249-448-59

Query Match 25.6%; Score 100; DB 3; Length 84;
Best Local Similarity 30.4%; Pred. No. 0.00097;
Matches 21; Conservative 8; Mismatches 24; Indels 16; Gaps 3;

QY 11 KCPSEIFSRCDGR-CQRFC-----PNVVPKPLCIKICAPGCVCLGYLRNKK 57
Db 5 QCGENEKYDSCGSKCEDKCKKYDGVBEEDDEEPNV---PCLVRVCHQDCVCEGFGYRNKD 61
QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

Search completed: February 26, 2003, 15:06:20
Job time : 32 secs

DB 5 QCGENEKYDSCGSKCEDKCKKYDGVBEEDDEEPNV---PCLVRVCHQDCVCEGFGYRNKD 61

QY 58 KVCVPRSKC 66
Db 62 DKCVSAEDC 70

RESULT 15
US-09-249-448-59
: Sequence 59, Application US/09249448
: Patent No. 6121435
: GENERAL INFORMATION:
: APPLICANT: Vlasuk, George Phillip
: APPLICANT: Stanssens, Patrick Eric Hugo
: APPLICANT: Messens, Joris Hilda Llieven
: APPLICANT: Lauwereys, Marc Josef
: APPLICANT: Laroche, Yves Rene
: APPLICANT: Jespers, Laurent Stephane
: APPLICANT: Gansemans, Yannick Georges Jozef
: APPLICANT: Moyle, Matthew
: APPLICANT: Bergum, Peter W.
: TITLE OF INVENTION: NEMATODE-EXTRACTED SERINE PROTEASE
: TITLE OF INVENTION: INHIBITORS AND ANTICOAGULANT
: TITLE OF INVENTION: PROTEIN
: NUMBER OF SEQUENCES: 356
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: STREET: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
: MEDIUM TYPE: storage
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: IBM P.C. DOS 5.0
: SOFTWARE: Word Perfect 5.1
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/249,448
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/809,455
: FILING DATE: April 17, 1997
: APPLICATION NUMBER: PCT/US95/13231
: FILING DATE: October 17, 1995
: APPLICATION NUMBER: 08/486,399
: FILING DATE: June 5, 1995
: APPLICATION NUMBER: 08/486,397
: FILING DATE: June 5, 1995
: APPLICATION NUMBER: 08/465,380
: FILING DATE: June 5, 1995
: APPLICATION NUMBER: 08/461,965
: FILING DATE: June 5, 1995
: APPLICATION NUMBER: 08/326,110
: FILING DATE: October 18, 1994
: ATTORNEY/AGENT INFORMATION:
: NAME: BIGGS, SUZANNE L.
: REGISTRATION NUMBER: 30,158
: REFERENCE/DOCKET NUMBER: 216/270
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: INFORMATION FOR SEQ ID NO: 59:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 84 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: ORIGINAL SOURCE:
: ORGANISM: Ancylostoma caninum

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:09:22 : Search time 17 Seconds
(without alignments)
378.883 Million cell updates/sec

Title: US-09-506-978-1
Perfect score: 391
Sequence: 1 GFGGGLGRGKCPSEIFSR.....CRLGLRNKKKVCVPRSKG 67

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 19399

Minimum DB seq length: 0
Maximum DB seq length: 67

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	80.5	20.6	63	2 S07127	chymotrypsin/elast
2	77.5	19.8	63	2 S08572	chymotrypsin/elast
3	70.5	18.0	62	2 S35098	trypsin inhibitor
4	69	17.6	60	2 S31723	metallothionein -
5	69	17.6	60	2 B27490	metallothionein B
6	68	17.4	60	2 S30567	metallothionein -
7	67	17.1	62	2 A35640	metallothionein 4
8	65	16.6	62	2 B35640	metallothionein IV
9	62	15.9	61	2 A37425	metallothionein 2
10	62	15.9	62	2 S54335	metallothionein-2c
11	62	15.9	62	2 S43336	metallothionein-2a
12	61.5	15.7	61	2 A33889	metallothionein 1
13	60.5	15.5	61	2 S3889	metallothionein 1A
14	60.5	15.5	61	1 SMHUIA	metallothionein A
15	60	15.3	52	2 S65712	metallothionein 1B
16	59.5	15.2	61	1 SMHUIB	metallothionein 1R
17	59.5	15.2	61	2 S69277	metallothionein 1R
18	59.5	15.2	63	2 S08190	metallothionein 1
19	59.5	15.2	63	2 A34905	metallothionein 1
20	59	15.1	60	2 JC2420	metallothionein -
21	59	15.1	61	1 SMBO2	metallothionein II
22	59	15.1	61	2 S00808	metallothionein Ia
23	59	15.1	61	2 S00809	metallothionein Ib
24	59	15.1	65	2 A38739	metallothionein -
25	58.5	15.0	61	1 SMWSI	metallothionein I
26	58.5	15.0	61	2 S47651	metallothionein 1H
27	58.5	15.0	61	2 I54574	metallothionein-1
28	58	14.8	61	2 S00810	metallothionein Ic
29	57.5	14.7	60	1 SMHUIA	metallothionein Ia

30	57.5	14.7	60	2 S38335	metallothionein -
31	57.5	14.7	61	1 SMHUIE	metallothionein 1E
32	57.5	14.7	61	1 SMHOB	metallothionein 1B
33	57.5	14.7	61	2 S47652	metallothionein 1X
34	56.5	14.5	61	1 SMHUI2	metallothionein 2
35	56.5	14.5	61	1 SMHUI2	metallothionein 2
36	56.5	14.5	61	1 SMHUIF	metallothionein 1F
37	56	14.3	64	2 A33825	metallothionein 1
38	55.5	14.2	61	1 SMRT1	metallothionein I
39	55	14.1	64	2 A25775	metallothionein A
40	54.5	13.9	48	2 G01478	thrombospondin-p50
41	54.5	13.9	61	1 SMHUI2C	metallothionein II
42	54.5	13.9	61	2 S00811	metallothionein II
43	54.5	13.9	61	2 B23689	metallothionein 2
44	54.5	13.9	61	2 I48173	metallothionein II
45	54.5	13.9	61	2 I48116	metallothionein II

ALIGNMENTS

RESULT 1

S07127
chymotrypsin/elastase inhibitor - common roundworm
C:Species: Ascaris lumbricoides (common roundworm)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C:Accession: S07127

R:Babin, D.R.; Peanasky, R.J.; Goos, S.M.
Arch. Biochem. Biophys. 232, 143-161, 1984
A:Title: The isoforms of chymotrypsin/elastase from Ascaris lumbricoides: the pr
A:Reference number: S07127; MUID:84255715; PMID:6564898
A:Accession: S07127
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-63 <BAB>
C:Superfamily: roundworm trypsin inhibitor

Query Match 20.6%; Score 80.5; DB 2; Length 63;
Best Local Similarity 30.3%; Pred. No. 0.23;
Matches 20; Conservative 10; Mismatches 23; Indels 13; Gaps 4;

QY	8	GRGKCPSEIFSRCDG---RC-----QRFCPNVVPKPLCIKICAPGCVRLGVRNKKKVC 60
DB	1	GOESGCPNEWTECTGCKCKGPDENTPCPLMCRPSC--ECSPG-----RGMRTNDGKC 54
QY	61	VPRSKC 66
DB	55	IPASQC 60

RESULT 2

S08572
chymotrypsin/elastase inhibitor - common roundworm
C:Species: Ascaris lumbricoides (common roundworm)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S08572

R:Babin, D.R.; Peanasky, R.J.; Goos, S.M.
Arch. Biochem. Biophys. 232, 143-161, 1984
A:Title: The isoforms of chymotrypsin/elastase from Ascaris lumbricoides: the pr
A:Reference number: S07127; MUID:84255715; PMID:6564898
A:Accession: S08572
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-63 <BAB>
C:Superfamily: roundworm trypsin inhibitor

Query Match 19.8%; Score 77.5; DB 2; Length 63;
Best Local Similarity 31.7%; Pred. No. 0.45;
Matches 19; Conservative 6; Mismatches 32; Indels 3; Gaps 2;

QY	9	RKCPSEIFSRCDGRCQRFCPNVVPKPLCIKICAPGCVCR--LGVRNKKKVCVPRSKC 66
DB	1	RKPCGKNEWTECTG--CELKCGQDENTPCALMCRPPSCSPGRMRTHDGKCVPVSEC 59

Copyright (c) 1993 - 2003 CompuGen Ltd.
GENCORE version 3.1.3

ein search, using sw model

' ; Search time 13 Seconds
(without alignments)
213.763 Million cell updates/sec

US-09-506-978-1

391
1 GGFGGLGGRGKCPSNEIFSR.....CRLGYLRNKKKVCVPRSKCG 67

BLOSUM62

Gapop 10.0 , Gapext 0.5

112892 seqs, 41476328 residues

hits satisfying chosen parameters. 6419

enath: 0

length: 67

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	104.5	26.7		56	1	AMCI_APIME	P56682 apis mellif
2	80.5	20.6		63	1	ICEL_ASCSU	P07851 ascaris suu
3	78.5	20.1		65	1	ICR2_ASCSU	P07852 ascaris suu
4	73	18.7		60	1	MT_PERFL	P52725 perca fluvi
5	70.5	18.0		62	1	ITRI_ASCSU	P19398 ascaris suu
6	69	17.6		60	1	MTB_ONCMY	P09862 oncorhynch
7	69	17.6		60	1	MT_ESOLU	P25127 esox lucius
8	68	17.4		60	1	MT_PLEPL	P07216 pleuronecte
9	68	17.4		60	1	MT_PSEAM	P55945 pseudopleur
10	67	17.1		62	1	MT1_HUMAN	P47944 homo sapien
11	66.5	17.0		61	1	MT_BALMY	O18842 balmo mys
12	66	16.9		60	1	MTB_SALSA	P52720 salmo salar
13	65	16.6		60	1	MT_ZOAVI	P52728 zoarces viv
14	65	16.6		62	1	MT4_MOUSE	P47945 mus musculu
15	63.5	16.2		60	1	MT_AMBWE	O42152 ambystoma m
16	62.5	16.0		62	1	MT4_CANFA	O9tui5 canis famili
17	62	15.9		62	1	MT2A_RABIT	P18055 oryctolagus
18	62	15.9		62	1	MT2C_RABIT	P80290 oryctolagus
19	61.5	15.7		61	1	MT1_BOVIN	P58280 bos taurus
20	60.5	15.5		61	1	MT1A_HUMAN	P04731 homo sapien
21	60.5	15.5		61	1	MT2_STECO	P14425 stenella co
22	60.5	15.5		61	1	MTA_ONCMY	P09861 oncorhynch
23	60	15.3		60	1	MT_PACMA	O91b50 pagrus majo
24	59.5	15.2		61	1	MT1B_HUMAN	P07438 homo sapien
25	59.5	15.2		61	1	MT1C_HUMAN	P093083 homo sapien
26	59.5	15.2		61	1	MT2_CANFA	Q9xst5 canis fami
27	59.5	15.2		62	1	MT2_CAEEL	P17512 caenorhabdi
28	59.5	15.2		63	1	MT1_COLL1	P15786 columba liv
29	59	15.1		60	1	MTB_CHAAC	P52724 chaenoceph
30	59	15.1		60	1	MTB_CHIHA	O13259 chionodrac
31	59	15.1		60	1	MTB_DICLA	Q9ptg9 dicentrarch
32	59	15.1		60	1	MTB_PAGBE	Q92145 pagothenia
33	59	15.1		60	1	MT_LIZAU	O13257 liza aurata

FT METAL 35 35 CLUSTER A.
 FT METAL 36 36 CLUSTER A.
 FT METAL 40 40 CLUSTER A.
 FT METAL 43 43 CLUSTER A.
 FT METAL 47 47 CLUSTER A.
 FT METAL 49 49 CLUSTER A.
 FT METAL 54 54 CLUSTER A.
 FT METAL 58 58 CLUSTER A.
 FT METAL 59 59 CLUSTER A.
 SQ SEQUENCE 60 AA; 5926 MW; E862A5EC16D7348A CRC64;

Query Match 18.7%; Score 73; DB 1; Length 60;
 Best Local Similarity 35.0%; Pred. No. 0.17;
 Matches 21; Conservative 3; Mismatches 22; Indels 14; Gaps 3;

OY 1 GGTGGLGGRGRCQRCPCPNVVKPLKIKICAPCCVRLGLYLNKKKVC 60
 DB 9 GGTGCGGSCYC-TNCSTTCKKSCCPGCGPCP-----CAGGVC-----KGKTC 54

RESULT 5
 ITRI_ASCSU STANDARD; PRT; 62 AA.
 AC P19398;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Trypsin inhibitor (ATI).
 OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
 OC Ascarididae; Ascaris.
 OX NCBI_TaxId=6253;
 RN [1]
 RP SEQUENCE.
 RA Peanasky R.J., Martzen M.R., Homandberg G.A., Cash J.M., Babin D.R.,
 RA Litwiller R.D.;
 RL (In) McInnis A.J. (eds.);
 RL Paradigms for eradicating helminthic parasites, pp.349-366,
 RL Alan R. Liss, New York (1987).
 RN [2]
 RP STRUCTURE BY NMR.
 RX MEDLINE=90212567; PubMed=2322539;
 RX Gronenborn A.M., Nilges M., Peanasky R.J., Clore G.M.;
 RT "Sequential resonance assignment and secondary structure
 RT determination of the Ascaris trypsin inhibitor, a member of a novel
 RT class of proteinase inhibitors.";
 RL Biochemistry 29:183-189(1990).
 CC -!- SIMILARITY: BELONGS TO THE ASCARIS TRYPSIN INHIBITOR FAMILY.
 DR PIR; S35098; S35098.
 DR PDB; LATA; 31-AUG-94.
 DR PDB; LATB; 31-AUG-94.
 DR PDB; LATD; 31-AUG-94.
 DR PDB; LATE; 31-AUG-94.
 DR InterPro; IPR002919; TIL_Cysrich.
 DR Pfam; PF01826; TIL; 1.
 KW Serine protease inhibitor; 3D-structure.
 FT DISULFID 5 38
 FT DISULFID 15 33
 FT DISULFID 18 29
 FT DISULFID 22 60
 FT DISULFID 40 54
 SQ SEQUENCE 62 AA; 6798 MW; 6438D25D2F76D3E1 CRC64;

Query Match 18.0%; Score 70.5; DB 1; Length 62;
 Best Local Similarity 29.5%; Pred. No. 0.33;
 Matches 18; Conservative 12; Mismatches 22; Indels 9; Gaps 5;

OY 11 KCPSS-NEIFSRCDGRCQRCPCPNVVKPLKIKICAP---GCVCRLGLYLNKKKVCVPRSK 65
 DB 4 KCTKPNQWTKCGG-CEGTCAQKIVP---CTRCCKPRCECIASAGFVRDAQNCIKFED 59
 OY 66 C 66

Db 60 C 60

RESULT 6
 MTB_ONCMY STANDARD; PRT; 60 AA.
 AC P09862;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metallothionein B (MT-B).
 GN MTB.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri), and
 OS Salvelinus alpinus (Arctic char).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxId=8022, 8036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O.mykiss;
 RX MEDLINE=88111026; PubMed=2448099;
 RA Bonham K., Zafarullah M., Gedamu L.;
 RT "The rainbow trout metallothioneins: molecular cloning and
 RT characterization of two distinct cDNA sequences.";
 RL DNA 6:519-528(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O.mykiss;
 RX MEDLINE=89039876; PubMed=3185557;
 RA Zafarullah M., Bonham K., Gedamu L.;
 RT "Structure of the rainbow trout metallothionein B gene and
 RT characterization of its metal-responsive region.";
 RL Mol. Cell. Biol. 8:4469-4476(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=O.mykiss; TISSUE=Liver;
 RX MEDLINE=91316146; PubMed=1859844;
 RA Kille P., Stephens P.E., Kay J.;
 RT "Elucidation of cDNA sequences for metallothioneins from rainbow
 RT trout, stone loach and pike liver using the polymerase chain
 RT reaction.";
 RL Biochim. Biophys. Acta 1089:407-410(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.alpinus; TISSUE=Liver;
 RA Gerpe M., Kling P., Olsson P.E.;
 RT "Metallothionein cDNA sequences and gene expression in arctic char
 RT (Salvelinus alpinus) following metal and PCB exposure.";
 RL Mar. Environ. Res. 46:551-554(1998).
 CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
 CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
 CC -!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
 CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
 CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
 CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
 CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
 CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M18104; AAA49566.1; -
 DR EMBL; M22487; AAA49567.1; -
 DR EMBL; X59394; CAA42037.1; -
 DR EMBL; AF013801; AAB66343.1; -
 DR PIR; A30818; A30818.
 DR PIR; B27490; B27490.

```

DR PIR: S16997; S16997.
DR PIR: S15509; S15509.
DR HSSP; P28184; 1J19.
DR InterPro; IPR003019; Metallthion.
DR Pfam; PF00131; metalthio; 1.
DR PRINTS; PS00860; MTVERTEBRATE.
DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW Metal-binding; Metal-thiolate cluster.
FT DOMAIN 1 28
FT METAL 29 60
FT METAL 4 4
FT METAL 6 6
FT METAL 12 12
FT METAL 14 14
FT METAL 18 18
FT METAL 20 20
FT METAL 23 23
FT METAL 25 25
FT METAL 28 28
FT METAL 32 32
FT METAL 33 33
FT METAL 35 35
FT METAL 36 36
FT METAL 40 40
FT METAL 43 43
FT METAL 47 47
FT METAL 49 49
FT METAL 54 54
FT METAL 58 58
FT METAL 59 59
FT METAL 60 60
SQ SEQUENCE 60 AA; 6033 MW; 9EA1E70FBES9B4EE CRC64;

Query Match 17.6%; Score 69; DB 1; Length 60;
Best Local Similarity 37.0%; Pred. No. 0.46; Indels 14; Gaps 3;
Matches 20; Conservative 2; Mismatches 18;

QY 7 GGRGKCPSEIFSRCDGRCPNVVPKPLCIKICAPGCVCRGLRNKKKVC 60
DB 15 GGSCKC-SNCACTSKCKSCCPCCPSCKS-----CASGCV-----KGKTC 54

RESULT 7
MT_ESOLU
ID MT_ESOLU STANDARD; PRT; 60 AA.
AC P25127;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein (MT).
GN MT.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Esociformes;
OC Esocidae; Esoc.
OX NCBI_TaxID=8010;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91316146; PubMed=1859844;
RA Kille P., Stephens P.E., Kay J.;
RT "Elucidation of cDNA sequences for metallothioneins from rainbow
RT trout, stone loach and pike liver using the polymerase chain
RT reaction.";
RL Biochim. Biophys. Acta 1216:55-64(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94032489; PubMed=8218416;
RA Kille P., Kay J., Sweeney G.E.;
RT "Analysis of regulatory elements flanking metallothionein genes in
RT Cd-tolerant fish (pike and stone loach).";
RL Biochim. Biophys. Acta 1216:55-64(1993).
KW -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC

```

RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).

-!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS: FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11 CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.

-!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X59392; CAA42035.1; -

CC EMBL; X70042; CAA49636.1; -

CC PIR; S17175; S17175.

CC PIR; S15503; S15503.

CC PIR; S38334; S38334.

CC PIR; S31723; S31723.

CC HSSP; P02795; 1MHU.

CC InterPro; IPR003019; Metallthion.

CC InterPro; IPR000006; Metallthion_vert.

CC Pfam; PF00131; metalthio; 1.

CC PRINTS; PR00860; MTVERTEBRATE.

CC PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.

KW Metal-binding; Metal-thiolate cluster.

FT DOMAIN 1 28

FT METAL 29 60

FT METAL 4 4

FT METAL 6 6

FT METAL 12 12

FT METAL 14 14

FT METAL 18 18

FT METAL 20 20

FT METAL 23 23

FT METAL 25 25

FT METAL 28 28

FT METAL 32 32

FT METAL 33 33

FT METAL 35 35

FT METAL 36 36

FT METAL 40 40

FT METAL 43 43

FT METAL 47 47

FT METAL 49 49

FT METAL 54 54

FT METAL 58 58

FT METAL 59 59

FT METAL 60 60

SQ SEQUENCE 60 AA; 5979 MW; 9EA1E43F95F8D97E CRC64;

Query Match 17.6%; Score 69; DB 1; Length 60;

Best Local Similarity 35.2%; Pred. No. 0.46;

Matches 19; Conservative 3; Mismatches 18; Indels 14; Gaps 3;

QY 7 GGRGKCPSEIFSRCDGRCPNVVPKPLCIKICAPGCVCRGLRNKKKVC 60

DB 15 GGSCKC-SNCACTSKCKSCCPCCPSCKS-----CASGCV-----KGKTC 54

RESULT 8

MT_PLEPL

ID MT_PLEPL STANDARD; PRT; 60 AA.

AC P07216;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Metallothionein (MT).

GN MT.

OS Pleuronectes platessa (Plaice).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Db 11 GGICMGDNCKTTCNKCKTKRSCCP--PGCAK-CARGCICKGG---SDKSCCP 62

RESULT 11

MT_BALMY
ID MT_BALMY STANDARD; PRT; 61 AA.
AC O18842;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein-IV (MT-IV).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kaysen J., O'Hara T., Goodwin T., Linnehan R., Hammond T.,
Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RL "Induction of a new metallothionein isoform (MT-IV) occurs during
differentiation of stratified squamous epithelia";
Biochemistry 33:7250-7259(1994).
CC -1- FUNCTION: SEEMS TO BIND ZINC AND COPPER. COULD PLAY A SPECIAL ROLE
IN REGULATING ZINC METABOLISM DURING THE DIFFERENTIATION OF
STRATIFIED EPITHELIA.
CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U07807; AAA20232.1; -
CC HSP; P02795; 2MHU.
CC MIM; 606206; -
CC InterPro; IPR003019; Metallothion.
CC InterPro; IPR000006; Metallothion_vert.
CC Pfam; PF00131; metalthio; 1.
CC PRINTS; PR00860; MTVERTBERATE.
CC PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
CC Metal-binding; Metal-thiolate cluster; Zinc; Copper.
FT METAL 6 6
FT METAL 8 8
FT METAL 14 14
FT METAL 16 16
FT METAL 20 20
FT METAL 22 22
FT METAL 25 25
FT METAL 27 27
FT METAL 30 30
FT METAL 34 34
FT METAL 35 35
FT METAL 37 37
FT METAL 38 38
FT METAL 42 42
FT METAL 45 45
FT METAL 49 49
FT METAL 51 51
FT METAL 58 58
FT METAL 60 60
FT METAL 61 61
SQ SEQUENCE 62 AA; 6419 MW; 36157CBA17BF28CC CRC64;
Query Match 17.1%; Score 67; DB 1; Length 62;
Best Local Similarity 35.1%; Pred. No. 0.76; Indels 4;
Matches 20; Conservative 8; Mismatches 23; Gaps 4;
7 GGRGKCPSEIFSRCDGR-CQRECPNVVPRPLCKICAPGCVCRGLYLRNKKKVCVP 62

RESULT 10

MT4_HUMAN
ID MT4_HUMAN STANDARD; PRT; 62 AA.
AC P47944;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein-IV (MT-IV).
GN MT4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-94271779; PubMed-8003488;
RA Qualif C.J., Findley S.D., Erickson J.C., Froelick G.J.,
Kelly E.J., Zambrowicz B.P., Palminter R.D.;
"Induction of a new metallothionein isoform (MT-IV) occurs during
differentiation of stratified squamous epithelia";
Biochemistry 33:7250-7259(1994).
CC -1- FUNCTION: SEEMS TO BIND ZINC AND COPPER. COULD PLAY A SPECIAL ROLE
IN REGULATING ZINC METABOLISM DURING THE DIFFERENTIATION OF
STRATIFIED EPITHELIA.
CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U07807; AAA20232.1; -
CC HSP; P02795; 2MHU.
CC MIM; 606206; -
CC InterPro; IPR003019; Metallothion.
CC InterPro; IPR000006; Metallothion_vert.
CC Pfam; PF00131; metalthio; 1.
CC PRINTS; PR00860; MTVERTBERATE.
CC PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
CC Metal-binding; Metal-thiolate cluster; Zinc; Copper.
FT METAL 6 6
FT METAL 8 8
FT METAL 14 14
FT METAL 16 16
FT METAL 20 20
FT METAL 22 22
FT METAL 25 25
FT METAL 27 27
FT METAL 30 30
FT METAL 34 34
FT METAL 35 35
FT METAL 37 37
FT METAL 38 38
FT METAL 42 42
FT METAL 45 45
FT METAL 49 49
FT METAL 51 51
FT METAL 58 58
FT METAL 60 60
FT METAL 61 61
SQ SEQUENCE 62 AA; 6419 MW; 36157CBA17BF28CC CRC64;
Query Match 17.1%; Score 67; DB 1; Length 62;
Best Local Similarity 35.1%; Pred. No. 0.76; Indels 4;
Matches 20; Conservative 8; Mismatches 23; Gaps 4;
7 GGRGKCPSEIFSRCDGR-CQRECPNVVPRPLCKICAPGCVCRGLYLRNKKKVCVP 62


```

FT METAL 40 40 CLUSTER A.
FT METAL 43 43 CLUSTER A.
FT METAL 47 47 CLUSTER A.
FT METAL 49 49 CLUSTER A.
FT METAL 54 54 CLUSTER A.
FT METAL 58 58 CLUSTER A.
FT METAL 59 59 CLUSTER A.
SQ SEQUENCE 60 AA; 6047 MW; E87B8D4FB2BC471A CRC64;

Query Match 16.6%; Score 65; DB 1; Length 60;
Best Local Similarity 37.08; Pred. No. 1.2; Indels 14; Gaps 4;
Matches 20; Conservative 3; Mismatches 17;

QY 7 GGRGKCPSPNEIFSRCDGRQFCFNVVPLKICAPGCVRLGRLNKKKVC 60
DB 15 GGSCKC-TNCSCTTCKRSCCPGSG-----CTK-CASGCVG-----KGKTC 54

RESULT 14
MT4_MOUSE
ID MT4_MOUSE STANDARD; PRT; 62 AA.
AC P47945;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein-IV (MT-IV).
GN MT4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6;
RX MEDLINE=94271779; PubMed=8003488;
RA Quafe C.J., Findley S.D., Erickson J.C., Froelick G.J.,
RA Kelly E.J., Zambowicz B.P., Palmer R.D.;
RT "Induction of a new metallothionein isoform (MT-IV) occurs during
RT differentiation of stratified squamous epithelia.";
RL Biochemistry 33:7250-7259(1994).
CC -!- FUNCTION: SEEMS TO BIND ZINC AND COPPER. COULD PLAY A SPECIAL ROLE
CC IN REGULATING ZINC METABOLISM DURING THE DIFFERENTIATION OF
CC STRATIFIED EPITHELIA.
CC -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN STRATIFIED SQUAMOUS
CC EPITHELIA ASSOCIATED WITH ORAL EPITHELIA, OESOPHAGUS, UPPER
CC STOMACH, TAIL, FOOTPADS, AND NEONATAL SKIN.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; U07808; AAA20233.1; -
CC HSSP; P18055; 1MRB.
CC MGD; MGI:99692; Mt4.
CC InterPro; IPR003019; Metallothion.
CC InterPro; IPR000006; Metallothion_vert.
CC Pfam; PF00131; metalthio; 1.
CC PRINTS; PR00860; MIVERTHEBRATE.
CC PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
CC Metal-binding: Metal-thiolate cluster; zinc; Copper.
FT METAL 6 6 CLUSTER B (BY SIMILARITY).
FT METAL 8 8 CLUSTER B (BY SIMILARITY).
FT METAL 14 14 CLUSTER B (BY SIMILARITY).
FT METAL 16 16 CLUSTER B (BY SIMILARITY).
FT METAL 20 20 CLUSTER B (BY SIMILARITY).
FT METAL 22 22 CLUSTER B (BY SIMILARITY).
FT METAL 25 25 CLUSTER B (BY SIMILARITY).
FT METAL 27 27 CLUSTER B (BY SIMILARITY).

```

```

FT METAL 30 30 CLUSTER B (BY SIMILARITY).
FT METAL 34 34 CLUSTER A (BY SIMILARITY).
FT METAL 35 35 CLUSTER A (BY SIMILARITY).
FT METAL 37 37 CLUSTER A (BY SIMILARITY).
FT METAL 38 38 CLUSTER A (BY SIMILARITY).
FT METAL 42 42 CLUSTER A (BY SIMILARITY).
FT METAL 45 45 CLUSTER A (BY SIMILARITY).
FT METAL 49 49 CLUSTER A (BY SIMILARITY).
FT METAL 51 51 CLUSTER A (BY SIMILARITY).
FT METAL 58 58 CLUSTER A (BY SIMILARITY).
FT METAL 60 60 CLUSTER A (BY SIMILARITY).
FT METAL 61 61 CLUSTER A (BY SIMILARITY).
SQ SEQUENCE 62 AA; 6276 MW; 8F78DA6D0744333A CRC64;

Query Match 16.6%; Score 65; DB 1; Length 62;
Best Local Similarity 33.9%; Pred. No. 1.2; Indels 10; Gaps 4;
Matches 21; Conservative 4; Mismatches 27;

QY 1 GGFGGGGLGGRGKCPSPNEIFSRCDGRQFCFNVVPLKICAPGCVRLGRLNKKKVC 60
DB 11 GGICICGDNCKCTWCSC-KTCRKSQPCPCP-----PGCAK-CARGCICKGG---SDKGCSC 60

QY 61 VP 62
DB 61 CP 62

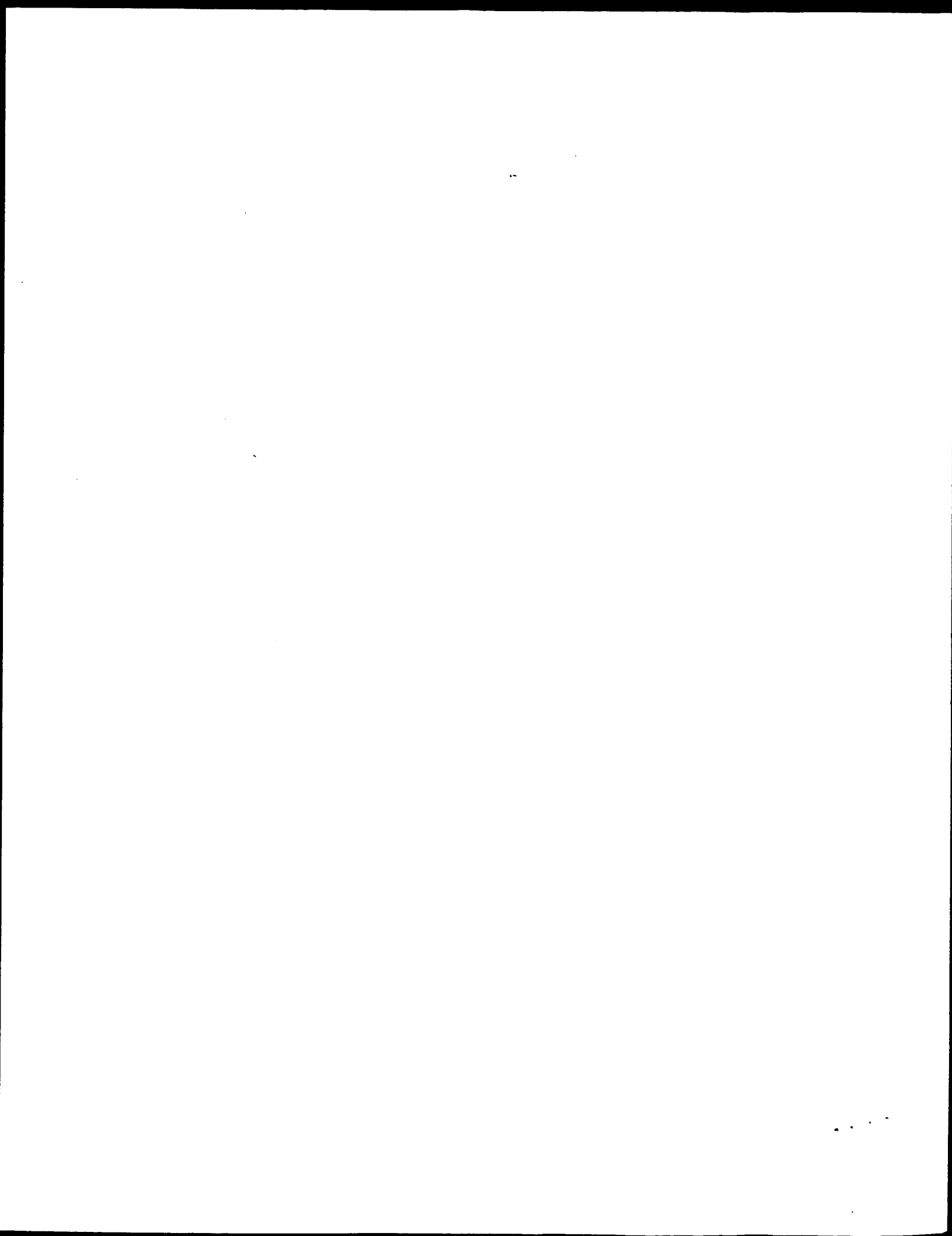
RESULT 15
MT_AMBME
ID MT_AMBME STANDARD; PRT; 60 AA.
AC Q42152;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein (MT) (AnMT).
GN MT-A.
OS Ambystoma mexicanum (Axolotl).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroides; Ambystomatidae;
OC Ambystoma.
OX NCBI_TaxID=8296;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=98127743; PubMed=9468225;
RA Saint-Jacques E., Guay J., Wirtanen L., Huard V., Stewart G.,
RA Seguin C.;
RT "Cloning of a complementary DNA encoding an Ambystoma mexicanum
RT metallothionein, AnMT, and expression of the gene during early
RT development.";
RL DNA Cell Biol. 17:83-91(1998).
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
CC -!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AF008583; AAB71835.1; -
CC HSSP; P02795; 1MHU.
CC InterPro; IPR003019; Metallothion.
CC InterPro; IPR000006; Metallothion_vert.
CC Pfam; PF00131; metalthio; 1.

```

Wed Feb 26 15:20:42 2003

DR	PRINTS; PRO0860; MTVERTEBRATE.
PR	PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW	Metal-binding; Metal-thiolate cluster.
FT	DOMAIN 1 28 BETA.
FT	DOMAIN 29 60 ALPHA.
FT	METAL 3 3 CLUSTER B.
FT	METAL 5 5 CLUSTER B.
FT	METAL 11 11 CLUSTER B.
FT	METAL 13 13 CLUSTER B.
FT	METAL 17 17 CLUSTER B.
FT	METAL 19 19 CLUSTER B.
FT	METAL 22 22 CLUSTER B.
FT	METAL 24 24 CLUSTER B.
FT	METAL 27 27 CLUSTER B.
FT	METAL 31 31 CLUSTER A.
FT	METAL 32 32 CLUSTER A.
FT	METAL 34 34 CLUSTER A.
FT	METAL 35 35 CLUSTER A.
FT	METAL 39 39 CLUSTER A.
FT	METAL 42 42 CLUSTER A.
FT	METAL 46 46 CLUSTER A.
FT	METAL 48 48 CLUSTER A.
FT	METAL 56 56 CLUSTER A.
FT	METAL 58 58 CLUSTER A.
FT	METAL 59 59 CLUSTER A.
SQ	SEQUENCE 60 AA; 5979 MW; AB05F4915953279 CRC64; .
 Query Match 16.2%; Score 63.5; DB 1; Length 60;	
Best Local Similarity 33.3%; Pred. No. 1.7; 23; Indels 7; Gaps 2;	
Matches 17; Conservative 4; Mismatches	
QY	1 GGFGGLGRGKCPNEIFSRCDGRQRCPNPVVPKPLCIKICAPGCVCRLG 51
Dh	8 GGSCSCAGSKCNCKTTS-CKKSCCSCCPCSECEK-----CGGGCVCKKG 51

Search completed: February 26, 2003, 15:10:59
Job time : 14 secs




```

RESULT 2
Q8WVB5 PRELIMINARY; PRT; 61 AA.
AC Q8WVB5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to metallothionein 1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC018190; AAH18190.1; -.
DR InterPro; IPR003019; Metallothion.
DR InterPro; IPR000006; Metallothion_vert.
DR Pfam; PF00131; metalthio; 1.
DR PRINTS; PR00860; MTEVERTEBRATE.
DR PROSITE; PS00203; METALLOTHIONEIN_VRT; UNKNOWN1.
SQ SEQUENCE 61 AA; 6034 MW; 3C4827F942E986B5 CRC64;

Query Match 16.2%; Score 63.5; DB 4; Length 61;
Best Local Similarity 34.8%; Pred. No. 0.47;
Matches 16; Conservative 6; Mismatches 19; Indels 5; Gaps 4;

QY 4 GGLGGRKCPSPNEIFSRCDGRQRCPCPNVVKPKLCKIKAPGCVCR 49
      | | | | | : | : | : | | | | | : |
DB 11 GSCACAGCKGNE--CKCTS--CKKSCGCPVG--CAK--CAQGGICK 51

RESULT 3
Q9PS78 PRELIMINARY; PRT; 51 AA.
AC Q9PS78;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Insulin-like growth factor binding protein-5, IGFBP-5 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96414636; PubMed=8817657;
RX Alexander S.V., Ehrenborg E., Luthman H., Powell D.R.;
RT "Conservation of IGFBP structure during evolution: cloning of chicken
RT insulin-like growth factor binding protein-5.";
RL Prog. Growth Factor Res. 6:159-165(1995).
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR Pfam; PF00219; IGFBP; 1.
DR SMART; SM00121; IB; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
FT NON_TER 1
FT NON_TER 51
SQ SEQUENCE 51 AA; 5127 MW; 662E453FACCB267 CRC64;

Query Match 16.1%; Score 63; DB 13; Length 51;
Best Local Similarity 45.2%; Pred. No. 0.46;
Matches 14; Conservative 2; Mismatches 11; Indels 4; Gaps 2;

QY 21 CRGRQRCPCPNVVKPKL-CIKICAPGCVCR 50
      |||: || ||| | : ||| |
DB 10 CDGRALSICP---PPPLGCELVRKPGCGCCL 37

RESULT 4
Q98TP9 PRELIMINARY; PRT; 40 AA.
AC Q98TP9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Metallothionein (MT) (Fragment).
GN MT.
OS Platichthys flesus (European flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neopteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidei; Pleuronectidae; Platichthys.
OX NCBI_TaxID=8260;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RA Williams T.D., Chipman J.K.;
RL "A DNA array to monitor the effects of environmental pollution on
RL European flounder (Platichthys flesus).";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
DR EMBL; AJ291833; CAC28138.1; -.
DR HSP; P02795; IMHU.
DR InterPro; IPR003019; Metallothion.
DR InterPro; IPR000006; Metallothion_vert.
DR Pfam; PF00131; metalthio; 1.
DR PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.
KW Chelation; Metal-binding; Metal-thiolate cluster.
FT NON_TER 1
FT NON_TER 40
SQ SEQUENCE 40 AA; 4010 MW; B5CEB05E4200855B CRC64;

Query Match 16.0%; Score 62.5; DB 13; Length 40;
Best Local Similarity 37.2%; Pred. No. 0.43;
Matches 16; Conservative 3; Mismatches 17; Indels 7; Gaps 2;

QY 7 GGRGKCPSPNEIFSRCDGRQRCPCPNVVKPKLCKIKAPGCVCR 49
      || | | | : | : | | | | | | | |
DB 5 GGSCTC-KNCSCTTCNKSCPCPCGCPK-----CASGCVCK 40

RESULT 5
Q99KF9 PRELIMINARY; PRT; 44 AA.
AC Q99KF9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 4.9 kDa protein.
GN Hypothetical 4.9 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004681; AAH04681.1; -.
DR MGD; MGI:1919666; 3110001K13Rik.
KW Hypothetical protein.
SQ SEQUENCE 44 AA; 4900 MW; C49523A82D448591 CRC64;

Query Match 16.0%; Score 62.5; DB 11; Length 44;
Best Local Similarity 33.3%; Pred. No. 0.47;
Matches 14; Conservative 5; Mismatches 14; Indels 9; Gaps 2;

QY 25 CQRCPCPNVVKPKLCKIKAPGCVCRIGLYLRNKKKVCVPRSKC 66
      | | | | | : | : | | | | | | | |

```


RESULT 14

```

Q8TDC4
ID Q8TDC4 PRELIMINARY; PRT; 61 AA.
AC Q8TDC4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Metallothionein 1Y.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Van Weyenbergh J., De Ley M.;
RT "Identification of molecular targets for zinc regulation of human
RL monocyte metabolism."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF479613; AAL85413.1; -
SQ SEQUENCE 61 AA; 6179 MW; 705F96E8B99D1CE CRC64;

Query Match 13.9%; Score 54.5; DB 4; Length 61;
Best Local Similarity 38.5%; Pred. No. 7.5;
Matches 15; Conservative 4; Mismatches 15; Indels 5; Gaps 4;

QY 10 GKCPSEIFSRGCRQFCPNVVPKPLCIKICAPGCV 48
Db | | | | | | | | | | | | | | | | | | | |
17 GSKCKE--CKCTS--CKKSCCPCPG--CAK--CAHGCVC 50

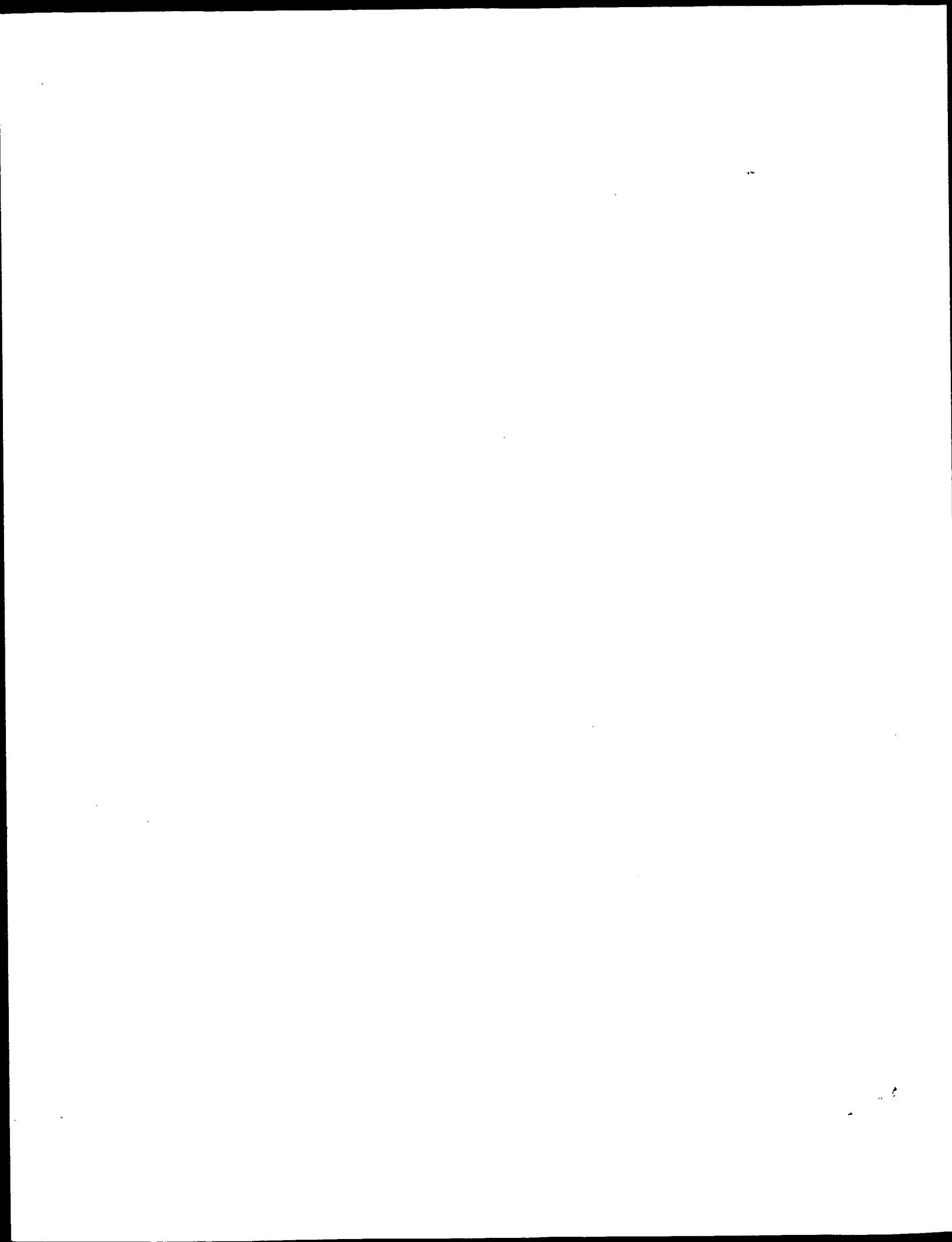
RESULT 15
Q9PVG7
ID Q9PVG7 PRELIMINARY; PRT; 39 AA.
AC Q9PVG7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Receptor tyrosine kinase Xmrk (Fragment).
GN XMRK.
OS Xiphophorus maculatus (Southern platyfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX NCBI_TaxID=8083;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SR;
RX MEDLINE=99126443; PubMed=9927468;
RA Gutbrod H., Scharl M.;
RT "Intragenic sex-chromosomal crossovers of Xmrk oncogene alleles affect
RT pigment pattern formation and the severity of melanoma in
RT Xiphophorus."
RL Genetics 151:773-783(1999).
DR EMBL: AF092693; AAD03714.1; -
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like; 1.
KW Kinase
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4195 MW; 42303BA05F6DB557 CRC64;

Query Match 13.6%; Score 53; DB 13; Length 39;
Best Local Similarity 42.9%; Pred. No. 7.7;
Matches 12; Conservative 3; Mismatches 9; Indels 4; Gaps 2;

QY 21 CDGRCQFCPNVVPKPL--CIKICAPGC 46
Db | | | | | | | | | | | | | | | | | | | |
5 CAEQCNRRCRG--PKPIDCCNEHCAGGC 30

```

Search completed: February 26, 2003, 15:12:38
 Job time : 92 secs



GenCore version 5.1.3
Copyright. (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:05:47 ; Search time 35 Seconds
(without alignments)
255.080 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGGGGLGGGRKCPSEIFSR.....CRLGLRNKKVCVPRKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 497633

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	391	100.0	67	ABB08330	Bee venom protein
2	82	21.0	61	AAV30434	Mature nematode ex
3	82	21.0	61	AAV30434	N. american nemat
4	68	17.4	58	AAV30433	Mature nematode ex
5	63.5	16.2	59	AAV57812	Trout metallothion
6	63.5	16.2	59	ABP32832	Human ORF1805 prot
7	62.5	16.0	62	ABP32832	Consensus sequence
8	62.5	16.0	63	AAO21321	Arabidopsis thalia
9	62	15.9	43	AAV57822	Human factor IX (h
10	62	15.9	61	AAV57822	Rabbit liver metal

11	61.5	15.7	67	22	AAV61455	Metallothionein do
12	61	15.6	47	20	AAV04648	Factor IXa catalyt
13	61	15.6	56	21	AAV44778	Human secreted pro
14	61	15.6	66	22	AAV45263	Propionibacterium
15	60	15.3	60	23	ABP01364	Human ORFX protein
16	59.5	15.2	61	22	AAU28133	Novel human secret
17	59	15.1	61	19	AAW61601	Human metallothion
18	58.5	15.0	39	20	AAV02082	KTX peptide used t
19	58	15.0	61	22	AAW40263	Human polypeptide
20	58	14.8	56	22	AAU48798	Propionibacterium
21	57.5	14.7	67	22	AAW79414	Human protein SEQ
22	57	14.6	10	21	AAV69218	N-terminal sequenc
23	57	14.6	40	5	AAV40692	Sequence encoded b
24	57	14.6	40	5	AAV40220	Partial sequence o
25	56.5	14.5	55	22	AAW39148	Human polypeptide
26	56.5	14.5	60	14	AAV40209	Sequence of human
27	56.5	14.5	60	21	AAV82332	Human metallothion
28	56.5	14.5	61	21	AAV12587	Human metallothion
29	56.5	14.5	61	21	AAV82331	Human metallothion
30	56.5	14.5	61	22	AAW78430	Human protein SEQ
31	56.5	14.5	61	23	ABP09810	Amino acid sequenc
32	56	14.3	66	23	ABP02424	Human ORFX protein
33	55.5	14.2	46	21	AAV12508	Beta-cellululin mute
34	55.5	14.2	47	21	AAV12507	Beta-cellululin mute
35	55.5	14.2	47	21	AAV12516	Beta-cellululin mute
36	55.5	14.2	48	21	AAV12515	Beta-cellululin mute
37	55.5	14.2	48	21	AAV12517	Beta-cellululin mute
38	55.5	14.2	48	21	AAV12518	Beta-cellululin mute
39	55.5	14.2	53	21	AAV12521	Beta-cellululin mute
40	55.5	14.2	60	21	AAW49082	Metallothionein fr
41	55.5	14.2	63	21	AAV57811	Chicken metallothi
42	55.5	14.2	67	23	ABP09671	Human ORFX protein
43	55	14.1	51	6	AAV50547	Protein analogue e
44	55	14.1	66	22	AAU55134	Propionibacterium
45	54.5	13.9	48	21	AAV12514	Beta-cellululin mute

ALIGNMENTS

RESULT 1

ABB08330
ID ABB08330 standard; protein; 67 AA.

XX ABB08330;

DT 18-JUN-2002 (first entry)

XX Bee venom protein Api m 6.01.

XX

KW Bee venom; isoform; immunosuppressant; vaccine; Api m 6; immune response;
KW Bee venom hypersensitivity; antibody; protein purification; Api m 6.01;
KW immunotherapy; allergen.

XX Apis sp.

XX WO200188085-A2.

XX 22-NOV-2001.

XX 16-FEB-2001; 2001WO-IB01736.

XX 18-FEB-2000; 2000US-0506978.

XX (ECOL-) ECOLE POLYTECHNIQUE FEDERALE LAUSANNE.

XX Spertini F;

XX WPI; 2002-082988/11.

PT New bee venom polypeptides, useful for modulating immune responses e.g.
PT in individual hypersensitive to the venom and for identifying
PT individual at risk for bee venom hypersensitivity

XX Example 2; Page 26; 32pp; English.
 CC The present sequence is that of one the four isoforms of Api m 6.
 CC designated Api m 6.01. The sequence represents the central amino acid
 CC sequence shared by all four isoforms (see ABB08331, ABB08332 and
 CC ABB08333). The specification describes a substantially pure polypeptide,
 CC Api m 6, derived from bee venom and found in four isoforms. The proteins
 CC of the invention have immunosuppressant activity and may form the basis
 CC of a vaccine. Api m 6 is useful for modulating an immune response, i.e.
 CC as an allergen for immunotherapy. The protein is useful for identifying
 CC an individual at risk for bee venom hypersensitivity. The method
 CC comprises administering Api m 6 to the individual and measuring an immune
 CC response raised, where a detectable immune response indicates that the
 CC individual is at risk for bee venom hypersensitivity. Antibodies specific
 CC for Api m 6 are useful for purifying the protein.

XX Sequence 67 AA;
 SQ Query Match 100.0%; Score 391; DB 23; Length 67;
 Best Local Similarity 100.0%; Pred. No. 8e-32;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GFGGLGGRGKCPNEIFSRCDGRCPNVVPKPLCIKICAPGCVCRGLYLRNKKVC 60
 Db 1 GFGGLGGRGKCPNEIFSRCDGRCPNVVPKPLCIKICAPGCVCRGLYLRNKKVC 60
 Qy 61 VPRSKCG 67
 Db 61 VPRSKCG 67

RESULT 2
 AAY30434
 ID AAY30434 standard; Protein; 61 AA.
 XX AC AAY30434;
 XX DT 15-NOV-1999 (first entry)
 XX DE Mature nematode extracted anticoagulant protein NamNAP.
 XX KW Nematode extracted anticoagulant protein; NAP; anticoagulant;
 KW serine protease inhibitor; NAP domain; factor VIIa/TF.
 XX OS Necator americanus.
 XX PN US5955294-A.
 XX PD 21-SEP-1999.
 XX PF 19-APR-1996; 96US-0634641.
 XX PR 19-APR-1996; 96US-0634641.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.
 PR 05-JUN-1995; 95US-0486397.
 PR 05-JUN-1995; 95US-0486399.
 PR 17-OCT-1995; 95WO-US13231.
 XX PA (CORV-) CORVAS INT INC.
 XX PI Bergum PW, Gansemans YGJ, Jespers LS, Laroche YR;
 PI Lauwereys MJ, Messens JHL, Moyle M, Stanssens PEH;
 PI Vlasuk GP;
 DR WPI; 1999-539569/45.
 XX PT Screening an isolated protein for Nematode-extracted Anticoagulant
 PT protein domains
 XX PT Disclosure; Columns 143-144; 197pp; English.

XX The present sequence represents a nematode extracted anticoagulant
 CC protein (NAP). The protein has activity as an anticoagulant and/or serine
 CC protease inhibitor. The protein contains at least one NAP domain which
 CC has selective inhibitory activity for factor VIIa/TF. The specification
 CC describes a method for screening an isolated protein at least one domain
 CC for factor VIIa/TF selective inhibitory activity. The method comprises
 CC determining the time to clotting effected by a concentration of the
 CC isolated protein in an ex vivo prothrombin time (PT) assay and an ex vivo
 CC activated partial thromboplastin time (aPTT) assay; calculating
 CC prolongation of clotting effected by the isolated protein in each of
 CC the PT and aPTT assay, with respect to a baseline clotting value for
 CC each assay, where prolongation of clotting is calculated as fold
 CC elevation of clotting time relative to a baseline clotting value, where
 CC a doubling of clotting time is deemed a two-fold elevation; and
 CC calculating a PT to aPTT prolongation ratio, where a ratio at least
 CC one is indicative of factor VIIa/TF inhibitory activity. The method is
 CC useful for determining if a protein has factor VIIa/TF inhibitory
 CC activity.

XX Sequence 61 AA;
 SQ Query Match 21.0%; Score 82; DB 20; Length 61;
 Best Local Similarity 30.9%; Pred. No. 0.37;
 Matches 17; Conservative 9; Mismatches 19; Indels 10; Gaps 2;
 Qy 12 CFSNEIFSRCDGRCPNVVPKPLCIKICAPGCV-----CRLGYLRNKKVCV 61
 Db 4 CPANEWRECGTPCEPKCNPMP-----DICTMNCIVDCQCKEGYKRHETKGCL 53

RESULT 3
 AAB15319
 ID AAB15319 standard; Protein; 61 AA.
 XX AC AAB15319;
 XX DT 19-DEC-2000 (first entry)
 XX DE N. americana nematode-extracted anticoagulant protein NamNAP.
 XX KW Nematode-extracted anticoagulant protein; NamNAP; blood clotting;
 KW canine hookworm; thrombosis; vaccine.
 XX OS Necator americanus.
 XX PN US6087487-A.
 XX PD 11-JUL-2000.
 XX PF 12-FEB-1999; 99US-0249451.
 XX PR 17-OCT-1995; 95WO-US13231.
 PR 17-APR-1997; 97US-0809455.
 PR 18-OCT-1994; 94US-0326110.
 PR 05-JUN-1995; 95US-0461965.
 PR 05-JUN-1995; 95US-0465380.
 PR 05-JUN-1995; 95US-0486397.
 PR 05-JUN-1995; 95US-0486399.
 XX PA (CORV-) CORVAS INT INC.
 XX PI Lauwereys MJ, Stanssens PEH, Jespers LS, Gansemans YGJ, Moyle M;
 PI Bergum PW, Messens JHL, LaRoche YR, Vlasuk GP;
 XX WPI; 2000-531359/48.
 DR N-PSDB; AAA73383.
 XX PT New cDNA molecule encoding a protein having factor Xa inhibitory
 PT activity for preventing and treating blood clotting disorders,
 PT comprises nematode-extracted anticoagulant protein domains
 XX PT Disclosure; Fig 16; 197pp; English.


```

XX SQ Sequence 59 AA:
Query Match 16.2%; Score 63.5; DB 21; Length 59;
Best Local Similarity 35.2%; Pred. No. 25;
Matches 19; Conservative 2; Mismatches 18; Indels 15; Gaps 3;

QY 7 GGRGKCPNEIFSRCDGRQRFQPCNVVPKPLCIKICAPGCVGLYLRNKKVC 60
Db 15 GSKCK-SNCACTSKKSCCPSPDCK-----CASGVC-----KGKTC 53

RESULT 6
ABP32832
ID ABP32832 standard; Protein; 59 AA.
AC
AC ABP32832;
DT
DT 09-JUL-2002 (first entry)
XX
XX Human ORF1805 protein, SEQ ID NO:3610.
XX
KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
KW disease monitoring; cytokine; cell proliferation; cell differentiation;
KW immune modulation; haematopoiesis regulation; tissue growth;
KW angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
KW behaviour; cancer; proliferative disorder; neurological disorder;
KW cardiovascular disease; immune system disorder; organ transplantation;
KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
KW hypothyroidism; cholesterol ester storage disease; infection; vulnery;
KW vasotropic; antipsoriatic; antidiabetic; cytostatic; nootropic;
KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
KW dermatological; analgesic; virucide; antibacterial; fungicide.
XX
OS Homo sapiens.
XX
XX WO200190366-A2.
XX
XX 29-NOV-2001.
XX
XX 24-MAY-2001; 2001WO-US17076.
XX
XX 24-MAY-2000; 2000US-206690P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Leach MD, Shimkets RA;
XX
XX WPI: 2002-106200/14.
XX
XX N-PSDB: ABN76859.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and disorders related to organ
XX transplantation.
XX
XX Claim 10; Page 1162; 2508pp; English.
XX
XX Sequences ABP31028-ABP35561 represent 4534 novel human proteins
XX designated ORF (open reading frame) 1-4534, and sequences ABN75054-
XX ABN79587 represent cDNAs encoding them. The invention also encompasses
XX polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
XX referred to as ORFX) proteins, polynucleotides at least 85% identical to
XX the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
XX polynucleotides, the recombinant production of ORFX proteins, antibodies
XX specific for ORFX proteins, methods of detecting ORFX polynucleotides and
XX polypeptides, methods of screening for modulators of ORFX expression or
XX activity, and methods of screening individuals for a predisposition to an
XX ORFX-associated disorder. The ORFX proteins of the invention have a wide
XX range of biological activities, such as cytokine, cell proliferation,
XX cell differentiation, immune modulation, haematopoiesis regulation,
XX

```

```

CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic activity, thrombolytic activity,
CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC and antifertile activity, and may also be involved in the determination
CC of bodily characteristics, fertility and behaviour. ORFX proteins,
CC nucleic acids and antibodies may be used in the treatment of cancers,
CC other proliferative disorders such as psoriasis and benign tumours,
CC neurological disorders such as epilepsy and Alzheimer's disease,
CC cardiovascular diseases, immune system disorders, disorders related to
CC organ transplantation, disorders of tissue growth and regeneration,
CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
CC storage disease, and infectious diseases caused by viral, bacterial,
CC fungal and other pathogens. ORFX nucleic acids may also be used as a
CC source of primers and probes, in the detection of ORFX genomic sequences
CC or transcripts, in the identification and cloning of homologous
CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
CC nucleic acids may additionally be used to produce transgenic animals
CC which may be useful for studying the function and/or activity of ORFX
CC protein, and in drug screening. The ORFX proteins may also be used as
CC immunogens to generate specific antibodies, which are useful in the
CC diagnosis, treatment and monitoring of ORFX-associated diseases.
XX
XX Sequence 59 AA;
SQ
Query Match 16.2%; Score 63.5; DB 23; Length 59;
Best Local Similarity 28.9%; Pred. No. 25;
Matches 13; Conservative 7; Mismatches 12; Indels 13; Gaps 2;

QY 25 QGRFCPNV---VPKPLCIKICAPGCVGLYLRNKKVCVPRSKC 66
Db 5 CAACVCPVCAAVCLPVCVQLVCMVC-----LCMHATMC 39

RESULT 7
ABP77804
ID ABP77804 standard; Protein; 62 AA.
XX
XX AC ABP77804;
XX
XX DT 27-SEP-2002 (first entry)
XX
XX Consensus sequence of mammalian metallothionein.
XX
XX Streptavidin; metallothionein; heavy metal ion; tumour cell imaging;
XX biotin; radiotherapy; mass spectrometry.
XX
XX Mammalia.
XX
XX Key Location/Qualifiers
XX Domain 1..29
XX FT /note= "beta domain"
XX Domain 30..62
XX FT /note= "alpha domain"
XX
XX US6391590-B1.
XX
XX 21-MAY-2002.
XX
XX 21-OCT-1991; 91US-0780717.
XX
XX 21-OCT-1991; 91US-0780717.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Sano T, Glazer AN, Cantor CR;
XX WPI: 2002-488386/52.
XX
XX A recombinant streptavidin-metallothionein chimeric protein is useful
XX to add or remove heavy metal ions into biotin-containing biological
XX material, particularly for tumour imaging, radiotherapy, and DNA or
XX protein labeling.
XX

```

PS Disclosure; Column 3; 9pp; English.

XX The present sequence represents mammalian metallothionein. It is
 CC used to produce chimeric proteins of the invention. The specification
 CC describes a recombinant bifunctional streptavidin-metallothionein
 CC chimeric protein. This protein is produced by introducing into a host
 CC cell nucleic acid encoding a bifunctional fusion protein having a
 CC streptavidin and a metallothionein moiety, and incubating the cell to
 CC express the fusion protein. The streptavidin moiety consists of residues
 CC 16-133 of mature streptavidin. The chimeric protein is used to
 CC incorporate heavy metal ions into biological materials containing biotin,
 CC or to remove heavy metal ions from the biological material. Specific uses
 CC include loading cancerous tissue with heavy metal ions for imaging of
 CC tumour cells and radiotherapy, and labeling DNA and proteins for
 CC detection on gels or blots by surface scanning mass spectrometry.

XX Sequence 62 AA;

Query Match 16.0%; Score 62.5; DB 23; Length 62;
 Best Local Similarity 37.3%; Pred. No. 32;
 Matches 19; Conservative 2; Mismatches 23; Indels 7; Gaps 3;

OY 1 GCGGLGGRGKCPSPNEIFSRCDGRQPCPNVVPKLCIKICAPGCVCRLG 51
 DB 10 GGSCTCAGSKCKCKCKCTS-CKKSCCCP-----VGCCK-CAQCGCVCKAG 53

RESULT 8

AA021321
 ID AA021321 standard; Protein; 63 AA.

XX AA021321;

XX 05-AUG-2002 (first entry)

XX Arabidopsis thaliana KCP-like protein, SEQ ID NO 76.

XX Antimicrobial; transgenic; plant; potato snakin antimicrobial protein;
 KW GAS44; GAS45; GAST1 homologue; lysine- and cysteine- rich peptide;
 KW KCP-like polypeptide; modulating; disease resistance.

XX Arabidopsis thaliana.

XX WO200222821-A2.

XX 21-MAR-2002.

XX 13-SEP-2001; 2001WO-US28429.

XX 13-SEP-2000; 2000US-232569P.

XX 11-SEP-2001; 2001US-0950933.

XX (PTON-) PIONEER HI-BRED INT INC.

XX Simmons CR, Navarro Acevedo PA;

XX WPI; 2002-425842/45.

XX New polynucleotide encoding lysine- and cysteine-rich peptides-like
 PT polypeptide useful for modulating the polypeptide level in a plant
 PT cell, enhancing disease resistance

XX Disclosure; Page 154; 163pp; English.

XX The invention relates to an isolated polynucleotide encoding a
 CC polypeptide which is related to potato snakin antimicrobial protein and
 CC GAS44 or GAS45 or GAST1 homologues, which is referred to as lysine- and
 CC cysteine- rich peptides (KCP)-like polypeptide, having a nucleotide
 CC sequence from 36 residues of defined base pairs, given in the
 CC specification. A recombinant expression cassette comprising the isolated
 CC polynucleotide of the invention is useful for modulating the level of
 CC (KCP)-like polypeptides in a plant cell, where the level of (KCP)-like
 CC polypeptides is increased, and disease resistance is enhanced. This

CC sequence represents a protein relating to the KCP-like proteins of the
 CC invention.

XX Sequence 63 AA;

Query Match 16.0%; Score 62.5; DB 23; Length 63;
 Best Local Similarity 37.8%; Pred. No. 33;
 Matches 17; Conservative 4; Mismatches 19; Indels 5; Gaps 3;

OY 20 RCDGRQRCPCPNV-VPKP---LCIKICAPGCVCRGLRNNKKVC 60
 DB 4 QCGGQCTRRCSNTYHKPCMFECQKCCAK-CLCVPPGTYGNKQVC 47

RESULT 9

AAB60286
 ID AAB60286 standard; Protein; 43 AA.

XX AAB60286;

XX 30-MAR-2001 (first entry)

XX Human factor IX (hFIX) exon-encoded fragment, SEQ ID NO:9.

XX Age-related gene regulation; liver-specific; gene expression;
 KW human factor IX; hFIX; AE5'; AE3'; age-regulatable expression construct;
 KW antisense therapy; gene therapy; thrombosis; cardiovascular disease;
 KW diabetes; Alzheimer's disease; Parkinson's disease; cancer; osteoporosis;
 KW osteoarthritis; dementia.

XX Homo sapiens.

XX WO200075279-A2.

XX 14-DEC-2000.

XX 06-JUN-2000; 2000WO-US15728.

XX 09-JUN-1999; 99US-0328925.

XX (UNMI) UNIV MICHIGAN.

XX Kurachi K, Kurachi S;

XX WPI; 2001-061708/07.

XX N-PSDB; AAF54018.

XX New regulatory elements that control age-related gene expression,
 PT useful in gene therapy and for reducing Factor IX expression

XX Disclosure; Fig 8C; 225pp; English.

XX The invention relates to nucleic acid sequences which regulate gene
 CC expression in an age-related manner and/or in a liver-specific manner.
 CC The invention identifies regions of the human factor IX (hFIX) gene, and
 CC a region of the human protein C (hPC) gene, which are age-related
 CC regulatory sequences. The hFIX age-related regulatory sequences are
 CC designated AE5' (AAF54016) and AE3' (AAF54017) and are found in the 5'
 CC UTR (at position 2164-2165 of AAF54018) and 3' UTR (at position
 CC 34383-35655 of AAF54018) respectively. These elements act synergistically
 CC to increase hFIX levels over the lifespan of an individual; however, they
 CC can independently exert effects on hFIX mRNA in an age-related manner,
 CC with AE5' acting to stabilise hFIX mRNA, and AE3' acting to increase hFIX
 CC mRNA levels, over time. AE5' also directs liver-specific expression. The
 CC hPC gene age-related regulatory sequence is found in the 5' UTR
 CC (AAF54081), and contains two PEA-3 (polyoma virus activator 3) elements
 CC 5'-GAGGAAGA-3' and 5'-CAGGAAG-3'. The age-related regulatory sequences of
 CC the invention, along with their homologues, variants and fragments, may
 CC be used in the construction of recombinant expression vectors for the
 CC expression of a desired sequence in an age-related fashion in a host
 CC cell. Preferred target genes for expression in such age-regulatable
 CC expression vectors include those encoding proteins involved in blood
 CC coagulation (e.g., the pro-coagulants factor IX and factor VIII, and the

CC anti-coagulants protein C and antithrombin III), human
 CC alpha-1-antitrypsin, PEA-3 protein and reporter proteins such as
 CC luciferase. Preferred promoters for use in such age-regulatable
 CC expression vectors include the human factor IX promoter, the T7 promoter,
 CC the T3 promoter and the SP6 promoter. The expression vectors of the
 CC invention may be used in gene therapy to provide age-related and/or
 CC liver-specific expression of target genes. Age-regulatable constructs may
 CC be used in the treatment of such age-related conditions such as
 CC thrombosis, cardiovascular disease, diabetes, Alzheimer's disease,
 CC Parkinson's disease, cancer, osteoporosis, osteoarthritis and dementia.
 CC Specifically, they may be used to express factor IX antisense mRNA in the
 CC treatment of thrombotic conditions associated with the natural
 CC age-related rise in factor IX expression. Transgenic cells or animals
 CC that contain vectors of the invention are useful as models of these
 CC diseases, in screening for potential therapeutic agents and for studying
 CC normal processes such as aging and gene expression. Fragments and
 CC homologues of age-related regulatory sequences, are useful as probes to
 CC detect, isolate or identify other such sequences in samples. The present
 CC sequence represents a fragment of hFIX.

XX Sequence 43 AA;

Query Match 15.9%; Score 62; DB 22; Length 43;
 Best Local Similarity 33.3%; Pred. No. 26;
 Matches 14; Conservative 5; Mismatches 15; Indels 8; Gaps 2;

QY 22 DGCRCQFCPNVVPKPLCIKICAPGCVRLGY-LRNKKKVCVP 62
 DB 8 NGRCEQFCNADNKVV-----CSCTEGYLAENQKSCP 42

RESULT 10

AAV57822

ID AAV57822 standard; protein; 61 AA.

XX AC AAV57822;

XX DT 22-MAR-2000 (first entry)

XX DE Rabbit liver metallothionein Class II amino acid sequence.

XX KW Metallothionein; metal recovery; remediation; heavy metal;
 XX KW precious metal; phytochelatin; green algae; Chlamydomonas reinhardtii.

XX OS Oryctolagus cuniculus.

XX PN WO9960838-A1.

XX PD 02-DEC-1999.

XX PF 28-MAY-1999; 99WO-US12007.

XX PR 28-MAY-1999; 98US-0087374.

XX PA (OHIS) UNIV OHIO STATE RES FOUND.

XX PI Sayre RT, Traina SJ;

XX DR WPI; 2000-086646/07.

XX PT Novel method for metal recovery, remediation and separation -

XX PS Disclosure; Fig 1; 86pp; English.

XX The present invention describes a transgenic algal cell (I) of the
 CC genus Chlamydomonas comprising reproductive genetic material comprising
 CC a nucleotide sequence capable of expressing chicken type I
 CC Metallothionein. Also described is a method of removing metal from
 CC an aqueous medium containing at least one dissolved or suspended
 CC metal. The transgenic algae are used for the selective separation of
 CC metals, particularly the separation of precious and desirable metals
 CC such as gold and uranium, from other metals such as cadmium, zinc and
 CC copper. The method can be used to facilitate the selective recovery of

CC precious and rare metals from mineral sources where aqueous media can
 CC be used, such as in natural surface water flows, ground water and where
 CC water may be introduced. The method is suitable for well-drilling,
 CC soil and water remediation arts, mining fields, and industrial
 CC engineering. The present sequence represents a Class II metallothionein
 CC given in the present invention.

XX Sequence 61 AA;

Query Match 15.9%; Score 62; DB 21; Length 61;
 Best Local Similarity 38.5%; Pred. No. 36;
 Matches 15; Conservative 2; Mismatches 10; Indels 12; Gaps 3;

QY 11 KCPSEIFSRCDGRQFCRPNVVPKPLCIKICAPGCVCR 49
 DB 25 KCTP-----CKKSCCSCP-----PGCAK-CAQGCICK 51

RESULT 11

AAB61455

ID AAB61455 standard; protein; 67 AA.

XX AC AAB61455;

XX DT 04-APR-2001 (first entry)

XX DE Metallothionein domain consensus.

XX KW TANGO; MANGO; colon; inflammation; tumor; renal; liver; lung;
 XX KW autoimmune; allergy; cardiovascular; brain; degenerative; placental;
 XX KW pancreatic; skeletal; muscle.

XX OS Synthetic.

XX PN WO200100672-A1.

XX PD 04-JAN-2001.

XX PF 29-JUN-2000; 2000WO-US18184.

XX PR 29-JUN-1999; 99US-0342687.

XX PA (MILL-) 'MILLENNIUM PHARM INC.

XX PI Holtzman DA, Barnes TM, Fraser CC, Sharp JD;

XX WPI; 2001-050127/06.

XX Isolated cDNAs encoding TANGO 244, TANGO 246, TANGO 275, TANGO 300 and
 CC MANGO 245 proteins, useful in the treatment of inflammatory diseases
 CC (e.g. idiopathic ulcerative colitis), tumors, renal disorders and liver
 CC disorders (e.g. jaundice) -

XX Disclosure; Fig 13; 262pp; English.

XX The present invention relates to cDNAs encoding TANGO 244,

CC TANGO 246, TANGO 275, TANGO 300 and MANGO 245 proteins.

CC The nucleic acids, proteins and protein modulators are useful for

CC treating colonic disorders, inflammatory diseases, tumors,

CC renal disorders, liver disorders, lung disorders, autoimmune diseases,

CC allergic diseases, cardiovascular diseases, brain disorders,

CC degenerative diseases placental, pancreatic, skeletal and muscle

XX disorders.

XX Sequence 67 AA;

Query Match 15.7%; Score 61.5; DB 22; Length 67;

Best Local Similarity 31.7%; Pred. No. 43;

Matches 19; Conservative 5; Mismatches 29; Indels 7; Gaps 2;

QY 1 GFGGLGRGKCPNENIFSRCDGRQFCRPNVVPKPLCIKICAPGCVRLGLENKKVC 60

DB 11 GGSCTGCTGCKCKNCKTS-CKKSCCSCCPAGCSK-----CAGGCVCCKGGAASETSKC 63

RESULT 12

AA04648
ID AAY04648 standard; peptide; 47 AA.

XX AC AAY04648;

XX XX 22-JUN-1999 (first entry)

XX DE Factor IXa catalytic and interacting domains.

XX KW Receptor; catalytic domain; Factor IXa; Factor Xa; tissue factor; angina;
KW blood clotting disorder; thrombosis; myocardial infarction;
KW reocluse; cerebrovascular disease; hypercoagulability; anticoagulant;
KW peripheral arterial occlusive disease; pulmonary embolism; cyclic.
XX OS Homo sapiens.

XX XX Key

XX FH Location/Qualifiers

XX FT Disulfide-bond 8..19

XX FT Domain 10..12

XX FT /note= "Factor VIIa interacting domain"

XX FT Disulfide-bond 13..29

XX FT Domain 14..16

XX FT /note= "catalytic domain"

XX FT Disulfide-bond 31..44

XX XX W09913062-A1.

XX PN 18-MAR-1999.

XX PD 08-SEP-1998; 98WO-GB02700.

XX PF 09-SEP-1997; 97GB-0019157.

XX PR (MATT/) MATTHEWS D P.

XX PA (NYCO-) NYCOMED IMAGING AS.

XX XX Fischer PM, Sakariassen KS;

XX PI WPI; 1999-215061/18.

XX DR New anticoagulant compounds

XX PT Disclosure; Fig 1; 50pp; English.

XX PS

XX XX

XX CC Peptides AAY04625-Y04647 represent claimed compounds which are capable

XX CC of interacting with an internal receptor in the catalytic domain of

XX CC Factor IXa (FIXa) or Factor X (FX) defined by the residues Ile290,

XX CC Ala291, Asp292, Tyr293, Thr294, Glu374 and Phe378 of FIXa, and Leu300,

XX CC Pro301, Glu302, Trp305, Ala306, Lys385 and Phe389 of FXa, or the ligand

XX CC defined by residues Cys95-Cys99 of FIXa or Cys96-Cys100 of FXa. The

XX CC compounds can be used to prevent the formation of a functional

XX CC FVIIa/FIXa or FVIIa/FXa complex, so can be used to combat or prevent

XX CC blood clotting disorders, e.g. thrombosis (particularly vascular

XX CC thrombosis or deep vein thrombosis), acute myocardial infarction,

XX CC restenosis, angina, reocluse, cerebrovascular disease, peripheral

XX CC arterial occlusive disease, hypercoagulability or pulmonary embolism.

XX CC They can also be used to prevent the occurrence of blood clotting

XX CC disorders caused by e.g. grafting surgery, vessel wall potency

XX CC restoration or sepsis.

XX XX

XX SQ Sequence 47 AA;

XX Query Match 15.6%; Score 61; DB 20; Length 47;

XX Best Local Similarity 34.1%; Pred. No. 36;

XX Matches 14; Conservative 4; Mismatches 15; Indels 8; Gaps 2;

QY 23 GRCQFPCNVVPKLCIKIPAGCVCVRLGY-LRNKKKVCVP 62

Db 13 GRCEQFCKNSADNKVV-----CSCTEGYLAENQKSCEP 46

RESULT 13

AA04778
ID AAB44778 standard; Protein; 56 AA.

XX AC AAB44778;

XX XX 12-FEB-2001 (first entry)

XX DE Human secreted protein sequence encoded by gene 17 SEQ ID NO:77.

XX KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; autoimmune disease; infection;
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;
KW cerebrovascular disorder; nervous system disorder; ocular disorder;
KW wound healing; skin aging; food additive; preservative.

XX OS Homo sapiens.

XX PN W0200058336-A1.

XX PD 05-OCT-2000.

XX PF 23-MAR-2000; 2000WO-US07726.

XX PR 26-MAR-1999; 99US-0126597.

XX PR 07-JAN-2000; 2000US-0174877.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX PA Rosen CA, Ruben SM, Komatsoulis G;

XX PI WPI: 2000-6023355/57.

XX DR N-PSDB; AAC79815.

XX XX Nucleic acid encoding human secreted proteins, used to treat, prevent,

XX PT ameliorate or diagnose medical conditions such as cancer, and

XX PT autoimmune diseases

XX PS Claim 11; Page 362; 391pp; English.

XX CC The polynucleotide sequences given in AAC79799 to AAC79848 encode the

XX CC human secreted proteins given in AAB44762 to AAB44811. AAB44812 to

XX CC AAB44829 represent human secreted polypeptide sequences and proteins

XX CC homologous to them, which are used in the exemplification of the present

XX CC invention. Human secreted proteins have activities based on the tissues

XX CC and cells the genes are expressed in. Examples of activities are:

XX CC immunosuppressive; antirheumatic; antirheumatic; antiproliferative;

XX CC cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic;

XX CC neuroprotective; antibacterial; virucide; fungicide; and

XX CC ophthalmological. The polynucleotides and polypeptides can be used to

XX CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,

XX CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used

XX CC in diagnosing a pathological condition or susceptibility to a

XX CC autoimmune disease, hyperproliferative disorders, cardiovascular

XX CC disorders, cerebrovascular disorders, angiogenesis, nervous system

XX CC disorders. The polypeptides can also be used to aid wound healing and

XX CC epithelial cell proliferation, to prevent skin aging due to sunburn, to

XX CC maintain organs before transplantation, for supporting cell culture of

XX CC primary tissues, to regenerate tissues and in chemotaxis. The

XX CC polypeptides can also be used as a food additive or preservative to

XX CC increase or decrease storage capabilities. AAC79799 to AAC79798 and

XX CC AAB44761 represent sequences used in the exemplification of the present

XX CC invention.

XX SQ Sequence 56 AA;

XX Query Match 15.6%; Score 61; DB 21; Length 56;

XX Best Local Similarity 29.8%; Pred. No. 42;

CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

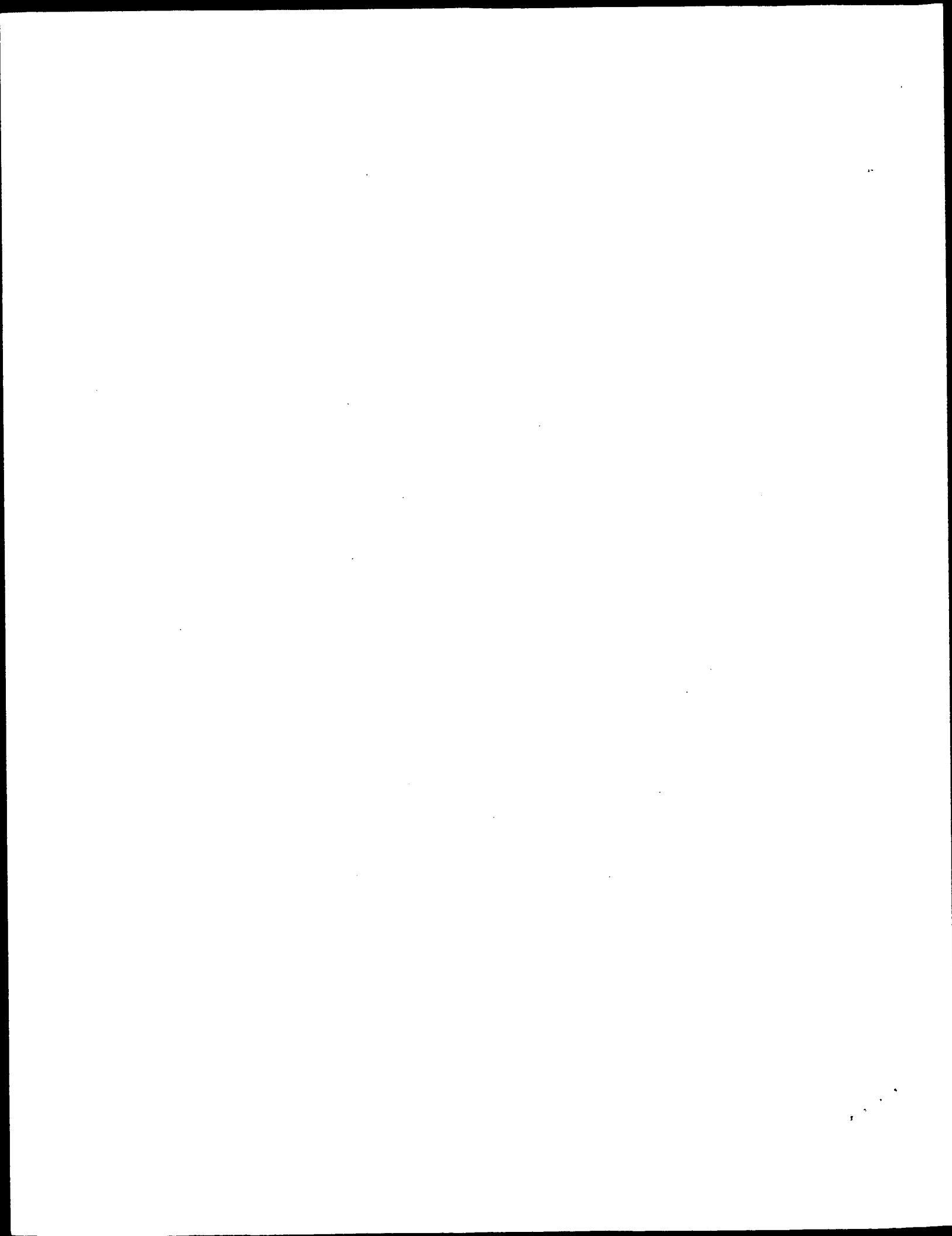
SQ Sequence 60 AA;

Query Match 15.3%; Score 60; DB 23; Length 60;
Best Local Similarity 35.5%; Pred. No. 56;
Matches 11; Conservative 3; Mismatches 17; Indels 0; Gaps 0;

Qy 36 PLCIKICAPGVCVCRGLGYLRNKKKVCVPRSKC 66

Db 3 PTCACVCVKPICACVCVCLLSXVCVCIPMCAC 33

Search completed: February 26, 2003, 15:10:41
Job time : 37 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:12:44 ; Search time 13 Seconds
(without alignments)
194.412 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGFGLGGRGKCPSEIFSR.....CRLGLRNKKKVCVPRSKG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 71782

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	16.1	60	10	US-09-790-264-59
2	62.5	16.0	63	9	US-09-950-933A-76
3	57.5	14.7	61	9	US-09-981-353-115
4	56.5	14.5	61	9	US-09-981-353-120
5	55.5	14.2	38	10	US-09-847-185-47
6	55.5	14.2	45	9	US-10-136-573A-11
7	55.5	14.2	45	9	US-09-877-665-11
8	55.5	14.2	45	9	US-10-215-862-11
9	55.5	14.2	45	10	US-09-817-647-11
10	55.5	14.2	46	9	US-10-201-945-12
11	54.5	13.9	58	10	US-09-865-578-11
12	53	13.6	46	10	US-09-864-761-34995
13	49	12.5	40	10	US-09-917-340-73
14	49	12.5	54	9	US-09-796-692-1773
15	49	12.5	64	9	US-09-796-692-1792
16	49	12.5	64	9	US-09-796-692-2056
17	48.5	12.4	51	10	US-09-864-761-37621
18	48.5	12.4	57	10	US-09-865-578-2
19	48.5	12.4	57	10	US-09-865-578-8

20 48.5 12.4 65 9 US-09-796-692-2465
21 48 12.3 40 10 US-09-917-340-68
22 48 12.3 51 10 US-09-854-864-6
23 48 12.3 64 10 US-09-864-761-47829
24 47.5 12.1 42 10 US-09-865-578-4
25 47 12.0 51 10 US-09-827-948-10
26 47 12.0 53 10 US-09-925-297-901
27 47 12.0 61 10 US-09-829-481-4
28 46.5 11.9 41 10 US-09-864-761-44588
29 46.5 11.9 50 9 US-10-138-158-10
30 46 11.8 37 10 US-09-829-481-11
31 46 11.8 37 10 US-09-917-340-57
32 46 11.8 50 9 US-09-932-172A-2
33 46 11.8 50 9 US-09-955-581-2
34 46 11.8 50 9 US-10-039-119A-2
35 46 11.8 50 9 US-10-138-158-2
36 46 11.8 50 9 US-10-138-158-11
37 46 11.8 50 9 US-10-138-158-12
38 46 11.8 50 10 US-09-299-473A-1
39 45.5 11.6 51 9 US-09-741-106-5
40 45.5 11.6 51 10 US-09-827-948-6
41 45.5 11.6 56 9 US-09-909-567B-52
42 45.5 11.6 63 9 US-09-950-933A-75
43 45.5 11.6 66 10 US-09-965-703-42
44 45 11.5 45 9 US-10-136-573A-13
45 45 11.5 45 9 US-09-877-665-13

ALIGNMENTS

RESULT 1
US-09-790-264-59
; Sequence 59, Application US/09790264
; Patent No. US20020028508A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNASTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-322001
; CURRENT APPLICATION NUMBER: US/09/790,264
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: US 09/065,661
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/298,531
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 09/065,363
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 09/337,930
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 09/102,705
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: US 09/363,630
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 09/124,538
; PRIOR FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 59
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-790-264-59

Query Match 16.1%; Score 63; DB 10; Length 60;
Best Local Similarity 38.1%; Pred. No. 7.7;
Matches 16; Conservative 2; Mismatches 20; Indels 4; Gaps 2;

QY 10 GKCPSEIFSRCDGRQRCFCPNVVPKPLCIKICAGCC--VCR 49
||||| | | | | | | | | | | | | | | | | |

Db 17 GKCPNP--PRSIGTCVELSCDQSCPNIQKCCSNGCGHVCK 56

RESULT 2

US-09-950-933A-76

; Sequence 76, Application US/09950933A

; Patent No. US20020166141A1

; GENERAL INFORMATION:

; APPLICANT: Simmons, Carl R.

; APPLICANT: Navarro, Pedro

; TITLE OF INVENTION: Antimicrobial Peptides and Methods of Use

; FILE REFERENCE: 35718/238472

; CURRENT APPLICATION NUMBER: US/09/950,933A

; CURRENT FILING DATE: 2001-09-11

; PRIOR APPLICATION NUMBER: 60/232,569

; PRIOR FILING DATE: 2000-09-13

; NUMBER OF SEQ ID NOS: 99

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 76

; LENGTH: 63

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

US-09-950-933A-76

Query Match 16.0%; Score 62.5; DB 9; Length 63;

Best Local Similarity 37.8%; Pred. No. 8.9;

Matches 17; Conservative 4; Mismatches 19; Indels 5; Gaps 3;

QY 20 RCDGRCQRCFPMV-VPKP---LCIKICAPGCVRLGYLELNKKVC 60

Db 4 QCGGQCTTRCSNTKYHKPCMFCQKCAK-CLCVPPGTGYNKQVC 47

RESULT 3

US-09-981-353-115

; Sequence 115, Application US/09981353

; Patent No. US20020160382A1

; GENERAL INFORMATION:

; APPLICANT: Lasek, Amy W.

; APPLICANT: Jones, David A.

; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER

; FILE REFERENCE: PA-0038 US

; CURRENT APPLICATION NUMBER: US/09/981,353

; CURRENT FILING DATE: 2001-10-11

; NUMBER OF SEQ ID NOS: 194

; SOFTWARE: PERL Program

; SEQ ID NO 115

; LENGTH: 61

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Incyte ID No. US20020160382A1 2685676CD1

US-09-981-353-115

Query Match 14.7%; Score 57.5; DB 9; Length 61;

Best Local Similarity 32.6%; Pred. No. 26;

Matches 15; Conservative 6; Mismatches 20; Indels 5; Gaps 4;

QY 4 GGLGGRGKCPSPNEIFSRCDGRCQRCFPMVVPKPLCIKICAPGCVCR 49

Db 11 GSCACAGSKCKE--CKCTS-CKKSCCSCPVG-CAK-CAQGCICK 51

RESULT 4

US-09-981-353-120

; Sequence 120, Application US/09981353

; Patent No. US20020160382A1

; GENERAL INFORMATION:

; APPLICANT: Lasek, Amy W.

; APPLICANT: Jones, David A.

; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER

; FILE REFERENCE: PA-0038 US

; CURRENT APPLICATION NUMBER: US/09/981,353

; CURRENT FILING DATE: 2001-10-11

; NUMBER OF SEQ ID NOS: 194

; SOFTWARE: PERL Program

; SEQ ID NO 120

; LENGTH: 61

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Incyte ID No. US20020160382A1 2757583CD1

US-09-981-353-120

Query Match 14.5%; Score 56.5; DB 9; Length 61;

Best Local Similarity 37.5%; Pred. No. 33;

Matches 15; Conservative 5; Mismatches 15; Indels 5; Gaps 4;

QY 10 GKCPSPNEIFSRCDGRCQRCFPMVVPKPLCIKICAPGCVCR 49

Db 17 GSKCKE--CKCTS-CKKSCCSCPVG-CSK-CAQGCICK 51

RESULT 5

US-09-847-185-47

; Sequence 47, Application US/09847185

; Patent No. US20020076392A1

; GENERAL INFORMATION:

; APPLICANT: Soo Hoo, William

; TITLE OF INVENTION: MEMBRANE-BOUND CYTOKINE COMPOSITIONS COMPRISING GM-CSF AND METHODS OF MODULATING AN IMMUNE RESPONSE USING SAME

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CAMPBELL & FLORES, LLP

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: United States

; ZIP: 92121

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/847,185

; FILING DATE: 01-May-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/201,931

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.

; REGISTRATION NUMBER: 31,815

; REFERENCE/DOCKET NUMBER: P-IM 2442

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619)535-9001

; TELEFAX: (619)535-8949

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 38 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 47:

US-09-847-185-47

Query Match 14.2%; Score 55.5; DB 10; Length 38;

Best Local Similarity 29.8%; Pred. No. 27;

Matches 14; Conservative 0; Mismatches 14; Indels 19; Gaps 2;

QY 2 GFGGLGGRGKCPSPNEIFSRCDGRCQRCFPMVVPKPLCIKICAPGCVCR 48

Db
1 GCGGAGGGGCC-----CTAGC-----ACCCACCGCTC 28

RESULT 6

```

US-10-136-573A-11
; Sequence 11, Application US/10136573A
; Patent No. US20020161200A1
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie Rose
; APPLICANT: Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related Ligands and
; TITLE OF INVENTION: Uses Therefor
; FILE REFERENCE: PI084R1C2
; CURRENT APPLICATION NUMBER: US/10/136,573A
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 09/480,977
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: US 08/999,437
; PRIOR FILING DATE: 1997-07-24
; PRIOR APPLICATION NUMBER: US 60/052,019
; PRIOR FILING DATE: 1997-07-09
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 11
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-136-573A-11

```

Query Match 14.2%; Score 55.5; DB 9; Length 45;
Best Local Similarity 35.9%; Pred. No. 31;
Matches 14; Conservative 5; Mismatches 13; Indels 7; Gaps 2;

18 FSRCDGRCQRFPCPNVWPPLCIKICA---PGCVCRIGYL 53
||||| : : : | : | : | ||| ||| :
2 FSRCPKQYKHYC---IKGRCFVVAEOTPCVCDEGI 36

RESULT 7

US-09-877-665-11
: Sequence 11, Application US/09877665
: Patent No. US20020164680A1
: GENERAL INFORMATION:
: APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
: TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
: Ligands and Uses Thereof
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 1 DNA Way
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WinPatIn (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/877,665
: FILING DATE: 08-Jun-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/09/109,206
: FILING DATE: 30-Jun-1998
: ATTORNEY/AGENT INFORMATION:
: NAME: Conley, Deirdre L.
: REGISTRATION NUMBER: 36,487
: REFERENCE/DOCKET NUMBER: P1084R1-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650/225-2066

TELEFAX: 650/952-9881
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 45 amino acids
 TYPE: Amino Acid
 TOPOLOGY: Linear
 FEATURE:
 NAME/KEY: hBTC.efg
 LOCATION: 1-45
 IDENTIFICATION METHOD:
 OTHER INFORMATION:
 SEQUENCE DESCRIPTION: SEQ ID NO: 11:
 US-09-877-665-11

Query Match 14.2%; Score 55.5; DB 9; Length 45;
Best Local Similarity 35.9%; Pred. No. 31;
Matches 14; Conservative 5; Mismatches 13; Indels

QY 18 FSRDGRGCRQRCFQPNVVPKPLCIKICA---PGCVRLCYL 53
||||| : : | | : | | ||| ||:
Db 2 FSRCPKQYKHVC----IKGCRFVVAEQTPTSCVCDGYI 36

RESULT 8

```

US-10-215-862-11
: Sequence 11, Application US/10215862
: Publication No. US20030036166A1
: GENERAL INFORMATION:
: APPLICANT: Godowski, Paul J.
: APPLICANT: Mark, Melanie Rose
: APPLICANT: Zhang, Dong Xiao
: TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related Ligands and
: TITLE OF INVENTION: Uses Thereof
: FILE REFERENCE: PI084R1D2C1
: CURRENT APPLICATION NUMBER: US/10/215,862
: CURRENT FILING DATE: 2002-09-24
: PRIOR APPLICATION NUMBER: US 09/126,663
: PRIOR FILING DATE: 1998-07-30
: PRIOR APPLICATION NUMBER: US 08/899,437
: PRIOR FILING DATE: 1997-07-24
: PRIOR APPLICATION NUMBER: US 60/052,019
: PRIOR FILING DATE: 1997-07-09
: NUMBER OF SEQ ID NOS: 23
: SEQ ID NO 11
: LENGTH: 45
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-215-862-11

```

Query Match 14.2%; Score 55.5; DB 9; Length 45;
Best Local Similarity 35.9%; Pred. No. 31;
Matches 14; Conservative 5; Mismatches 13; Indels

18 FSRCDGRQRCFCPVPVVKPLCIKICA---PGCVCRIGYL 53
||||| : : : | | : | | | | | :
2 FSRCPKQYKHYC----IKGRCFVVAEQTPSCVCDGYI 36
Ddb

RESULT 9

US-09-817-647-11
Sequence 11, Application US/09817647
Patent No. US20020082229A1
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
Ligands and Uses Therefor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:10:48 ; Search time 352 Seconds
(without alignments)
122.719 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGFGLGGRGKCPSEIFSR.....CRLGYLRNKKVCVPRSKG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 2759142

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PTCUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	391	100.0	67	19 US-09-506-978-1	Sequence 1, Appli
2	391	100.0	67	25 US-10-174-151-1	Sequence 1, Appli
3	391	100.0	67	26 US-10-204-145-1	Sequence 1, Appli
4	113.5	29.0	62	27 US-60-142-896-1491	Sequence 1491, Ap
5	113.5	29.0	62	27 US-60-145-138-701	Sequence 701, App
6	88.5	22.6	62	21 US-09-791-537-9664	Sequence 9664, Ap

7	87	22.3	39	27	US-60-160-203-3377	Sequence 3377, Ap
8	87	22.3	39	27	US-60-169-840-5102	Sequence 5102, Ap
9	86.5	22.1	54	27	US-60-162-247-3339	Sequence 3339, Ap
10	84.5	21.6	53	27	US-60-163-123-1354	Sequence 1354, Ap
11	84.5	21.6	53	27	US-60-163-123-1440	Sequence 1440, Ap
12	84	21.5	60	27	US-60-160-203-4074	Sequence 4074, Ap
13	84	21.5	60	27	US-60-163-123-1552	Sequence 1552, Ap
14	84	21.5	60	27	US-60-163-123-1677	Sequence 1677, Ap
15	82.5	21.1	63	27	US-60-169-840-5359	Sequence 5359, Ap
16	82	21.0	61	16	US-09-249-473-61	Sequence 61, Appl
17	82	21.0	61	18	US-09-498-272-61	Sequence 61, Appl
18	82	21.0	61	18	US-09-498-556-61	Sequence 61, Appl
19	80.5	20.6	61	21	US-09-791-537-13914	Sequence 13914, A
20	80.5	20.6	63	21	US-09-791-537-86288	Sequence 86288, A
21	78.5	20.1	65	21	US-09-791-537-86289	Sequence 86289, A
22	77.5	19.8	63	21	US-09-791-537-4553	Sequence 4553, Ap
23	76	19.4	52	27	US-60-162-243-691	Sequence 691, App
24	75.5	19.3	64	27	US-60-196-174-1132	Sequence 1132, Ap
25	73	18.7	43	27	US-60-164-762-848	Sequence 848, App
26	73	18.7	60	21	US-09-791-537-125018	Sequence 125018, A
27	72	18.4	43	27	US-60-164-762-1114	Sequence 1114, Ap
28	71.5	18.3	50	27	US-60-160-203-3591	Sequence 3591, Ap
29	71.5	18.3	53	27	US-60-160-203-3421	Sequence 3421, Ap
30	70.5	18.0	62	21	US-09-791-537-103884	Sequence 103884, A
31	70	17.9	53	27	US-60-170-373-3452	Sequence 3452, Ap
32	69.5	17.8	51	27	US-60-163-123-1638	Sequence 1638, Ap
33	69.5	17.8	56	27	US-60-188-162-2853	Sequence 2853, Ap
34	69	17.6	55	9	US-08-548-186-196	Sequence 196, App
35	69	17.6	60	21	US-09-791-537-137772	Sequence 137772, A
36	69	17.6	60	21	US-09-791-537-137774	Sequence 137774, A
37	68.5	17.5	52	21	US-09-791-537-90125	Sequence 90125, A
38	68	17.4	58	16	US-09-249-473-60	Sequence 60, Appl
39	68	17.4	58	18	US-09-498-272-60	Sequence 60, Appl
40	68	17.4	58	18	US-09-498-556-60	Sequence 60, Appl
41	68	17.4	60	21	US-09-791-537-31406	Sequence 31406, A
42	68	17.4	60	21	US-09-791-537-44038	Sequence 44038, A
43	67	17.1	62	21	US-09-791-537-51265	Sequence 51265, A
44	67	17.1	62	27	US-60-389-987-1874	Sequence 1874, Ap
45	67	17.1	62	27	US-60-412-418-1874	Sequence 1874, Ap

ALIGNMENTS

RESULT 1

US-09-506-978-1

; Sequence 1, Application US/09506978

; GENERAL INFORMATION:

; APPLICANT: Spertini, Francois

; TITLE OF INVENTION: NOVEL BEE VENOM POLYPEPTIDES AND METHODS OF USE THEREOF

; FILE REFERENCE: 18519-001

; CURRENT APPLICATION NUMBER: US/09/506,978

; CURRENT FILING DATE: 2000-02-18

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 67

; TYPE: PRT

; ORGANISM: Apis mellifera

US-09-506-978-1

Query Match 100.0%; Score 391; DB 19; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.le-32;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGLGGRGKCPSEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKVC 60
Db 1 GGFGLGGRGKCPSEIFSRCDGRQFCPNVVPKPLCIKICAPGCVRLGYLRNKKVC 60
QY 61 VPRSKG 67
Db 61 VPRSKG 67


```
RESULT 12
US-60-160-203-4074
; Sequence 4074, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4074
; LENGTH: 60
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-203-4074

Query Match      21.5%; Score 84; DB 27; Length 60;
Best Local Similarity 31.7%; Pred. No. 0.83;
Matches 20; Conservative 10; Mismatches 13; Indels 20; Gaps 4;

QY 11 KCPNNEIFSR-----DGRQRCPCNVVVKPLCIKICAPGCVCRGLGYLRNKKV 59
:|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
DB 7 ECPAYSSYTNCLPSCSPSCWDLGRCE---GAKVP-----SACAEGCICQPGYVLSDEK- 57

QY 60 CVP 62
| | |
DB 58 CVP 60

RESULT 13
US-60-163-123-1552
; Sequence 1552, Application US/60163123
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000137
; CURRENT APPLICATION NUMBER: US/60/163,123
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1986
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1552
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Human
US-60-163-123-1552

Query Match      21.5%; Score 84; DB 27; Length 60;
Best Local Similarity 31.7%; Pred. No. 0.83;
Matches 20; Conservative 10; Mismatches 13; Indels 20; Gaps 4;

QY 11 KCPNNEIFSR-----DGRQRCPCNVVVKPLCIKICAPGCVCRGLGYLRNKKV 59
:|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
DB 7 ECPAYSSYTNCLPSCSPSCWDLGRCE---GAKVP-----SACAEGCICQPGYVLSDEK- 57

QY 60 CVP 62
| | |
DB 58 CVP 60

RESULT 14
US-60-163-123-1677
; Sequence 1677, Application US/60163123
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
```

```
; FILE REFERENCE: CL000137
; CURRENT APPLICATION NUMBER: US/60/163,123
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1986
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1677
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Human
US-60-163-123-1677

Query Match      21.5%; Score 84; DB 27; Length 60;
Best Local Similarity 31.7%; Pred. No. 0.83;
Matches 20; Conservative 10; Mismatches 13; Indels 20; Gaps 4;

QY 11 KCPNNEIFSR-----DGRQRCPCNVVVKPLCIKICAPGCVCRGLGYLRNKKV 59
:|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
DB 7 ECPAYSSYTNCLPSCSPSCWDLGRCE---GAKVP-----SACAEGCICQPGYVLSDEK- 57

QY 60 CVP 62
| | |
DB 58 CVP 60

RESULT 15
US-60-169-840-5359
; Sequence 5359, Application US/60169840
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000164
; CURRENT APPLICATION NUMBER: US/60/169,840
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 9628
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5359
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Human
US-60-169-840-5359

Query Match      21.1%; Score 82.5; DB 27; Length 63;
Best Local Similarity 32.1%; Pred. No. 1.2;
Matches 18; Conservative 8; Mismatches 29; Indels 1; Gaps 1;

QY 12 CPSNEIFSRCDGRQRCPCNVVVKPLCIKICAPGCVCRGLGYLRNKKVKVPSKCG 67
:|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|: :|:
DB 1 CPAHSHYSICTRTCGSCAALSGLTGCTTRCFEGCBDDRELLS-QGVCIPVQDCG 55

Search completed: February 26, 2003, 15:19:21
Job time : 353 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:11:03 : Search time 24 seconds
(without alignments)
255.149 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GFGGLGGRGKCPNSIEFSR.....CRLGLRNKKKVCVPRSKG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 486122 seqs, 91396495 residues

Total number of hits satisfying chosen parameters: 231224

Minimum DB seq length: 0

Maximum DB seq length: 67

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match %	Length	DB ID	Description
1	119	30.4	67	6	US-10-038-854-185	Sequence 185, App
2	63	16.1	60	6	US-10-289-353-59	Sequence 59, Appl
3	62.5	16.0	62	6	US-10-285-876-5	Sequence 5, Appl
4	61	15.6	66	1	PCT-US02-32727-6458	Sequence 6458, App
5	61	15.6	66	5	US-09-978-825-6458	Sequence 6458, App
6	61	15.6	66	6	US-10-057-498-6458	Sequence 6458, App
7	58.5	15.0	61	6	US-10-170-385-239	Sequence 239, App
8	58	14.8	56	1	PCT-US02-32727-9993	Sequence 9993, App
9	58	14.8	56	5	US-09-978-825-9993	Sequence 9993, App
10	58	14.7	56	6	US-10-057-498-9993	Sequence 9993, App
11	57.5	14.7	61	6	US-10-170-385-241	Sequence 241, App
12	57.5	14.7	61	7	US-60-440-068-445	Sequence 445, App
13	57	14.6	67	5	US-09-724-676-81522	Sequence 81522, A
14	57	14.6	67	5	US-09-724-676A-81522	Sequence 81522, A
15	56.5	14.5	61	6	US-10-170-385-265	Sequence 265, App
16	56.5	14.5	61	7	US-60-423-586-136	Sequence 136, App
17	56.5	14.5	61	7	US-60-427-194-136	Sequence 136, App
18	55.5	14.2	46	5	US-09-857-815A-4	Sequence 4, Appl
19	55.5	14.2	47	5	US-09-857-815A-3	Sequence 3, Appl
20	55.5	14.2	47	5	US-09-857-815A-12	Sequence 12, Appl
21	55.5	14.2	48	5	US-09-857-815A-11	Sequence 11, Appl
22	55.5	14.2	48	5	US-09-857-815A-14	Sequence 14, Appl
23	55.5	14.2	49	5	US-09-857-815A-10	Sequence 10, Appl
24	55.5	14.2	49	5	US-09-857-815A-13	Sequence 13, Appl
25	55.5	14.2	58	6	US-10-231-778-229	Sequence 229, App
26	55	14.1	66	1	PCT-US02-32727-16329	Sequence 16329, A

ALIGNMENTS

RESULT 1

US-10-038-854-185

: Sequence 185, Application US/10038854

: GENERAL INFORMATION: Kimberly A

: APPLICANT: Spytek, Kimberly A

: APPLICANT: Li, Li

: APPLICANT: Wolenc, Adam R

: APPLICANT: Vernet, Corine

: APPLICANT: Eisen, Andrew J

: APPLICANT: Liu, Xiaohong

: APPLICANT: Malyankar, Uriel M

: APPLICANT: Shimkets, Richard A

: APPLICANT: Tchernev, Velizar

: APPLICANT: Spaderna, Steven K

: APPLICANT: Gorman, Linda

: APPLICANT: Kekuda, Ramesh

: APPLICANT: Patturajan, Meera

: APPLICANT: Gusev, Vladimir Y

: APPLICANT: Gangolli, Esha A

: APPLICANT: Guo, Xiaojia S

: APPLICANT: Shenoy, Suresh G

: APPLICANT: Rastelli, Luca

: APPLICANT: Casman, Stacie J

: APPLICANT: Boldog, Ferenc

: APPLICANT: Burgess, Catherine E

: APPLICANT: Edinger, Shlomit R

: APPLICANT: Ellerman, Karen

: APPLICANT: Gunther, Erik

: APPLICANT: Smithson, Glenda

: APPLICANT: Millet, Isabelle

: APPLICANT: MacDougall, John R

: TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

: FILE REFERENCE: 21402-230

: CURRENT APPLICATION NUMBER: US/10/038,854

: PRIOR FILING DATE: 2003-01-22

: PRIOR APPLICATION NUMBER: 60/258,928

: PRIOR FILING DATE: 2000-12-29

: PRIOR APPLICATION NUMBER: 60/259,415

: PRIOR FILING DATE: 2001-01-02

: PRIOR APPLICATION NUMBER: 60/259,785

: PRIOR FILING DATE: 2001-01-04

: PRIOR APPLICATION NUMBER: 60/269,814

: PRIOR FILING DATE: 2001-02-20

: PRIOR APPLICATION NUMBER: 60/279,832

: PRIOR FILING DATE: 2001-03-29

: PRIOR APPLICATION NUMBER: 60/279,833

: PRIOR FILING DATE: 2001-03-29

: PRIOR APPLICATION NUMBER: 60/279,863

: PRIOR FILING DATE: 2001-03-29

Sequence 16329, A
Sequence 16329, A
Sequence 6751, Ap
Sequence 6751, Ap
Sequence 6751, Ap
Sequence 44, Appl
Sequence 12069, A
Sequence 243, App
Sequence 81494, A
Sequence 81494, A
Sequence 20050, A
Sequence 20050, A
Sequence 20050, A
Sequence 33, Appl
Sequence 33, Appl
Sequence 233, App
Sequence 47, Appl
Sequence 4852, Ap
Sequence 4852, Ap

55 14.1 66 5 US-09-978-825-16329
55 14.1 66 6 US-10-057-498-16329
54.5 13.9 60 1 PCT-US02-32727-6751
54.5 13.9 60 5 US-09-978-825-6751
54.5 13.9 60 6 US-10-057-498-6751
54 13.8 53 5 US-09-857-815A-44
53 13.6 46 6 US-10-203-138A-12069
53 13.6 61 6 US-10-170-385-243
53 13.6 67 5 US-09-724-676-81494
53 13.6 67 5 US-09-724-676A-81494
52.5 13.4 57 1 PCT-US02-32727-20050
52.5 13.4 57 5 US-09-978-825-20050
52.5 13.4 57 6 US-10-057-498-20050
51 13.0 36 6 US-10-004-378A-83
51 13.0 61 6 US-10-218-140-3458
51 12.9 57 6 US-10-231-778-233
50 12.8 36 1 PCT-US02-18256-47
50 12.8 44 1 PCT-US02-32727-4852
50 12.8 58 5 US-09-978-825-4852

```

RESULT 3
US-10-285-876-5
; Sequence 5, Application US/10285876
; GENERAL INFORMATION:
; APPLICANT: Sano, Takeshi
;          Glazer, Alexander N
;          Cantor, Charles R
; TITLE OF INVENTION: Metallothionein Derivatives with
;                   Biological Recognition Specificity
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/285,876
; FILING DATE: 01-Nov-2002
; CLASSIFICATION: <Unknown>
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/07/780,717
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B91-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 62 amino acids
; TYPE: amino acid
; STRANDEDNESS: Not Relevant
; TOPOLOGY: Not Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-285-876-5

Query Match           16.0%; Score 62.5; DB 6; Length 62
Best Local Similarity 37.3%; Pred. No.11;
Matches 19; Conservative 2; Mismatches 23; Indels

QY      1 GGFGGLGGRGKCPSEIETSRCDGRCFPCPNVVPKPICIKICACGCVCRLG    51
        || | | | | | | | | | | | | | | | | | | | | | | | |
DB       10 GGSCTCAGSKCKECKCKTS-CKKSCCCSCP-----VGCAK-CAQGCVCKAG    53


RESULT 4
PCT-US02-32727-6458
; Sequence 6458, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shvian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
```

```

; APPLICANT: Douglass,John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 6458
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Propioni acnes
PCT-US02-32727-6458

```

```

Query Match      15.6%; Score 61; DB 1; Length 66;
Best Local Similarity 31.8%; Pred. No. 17;
Matches 21; Conservative 6; Mismatches 27; Indels 12; Gaps 4;

QY  2  GFGGLGGRGKC---PSNEIFSRCDGRQRCPCPNVVPKPLCIKICAPGCVRLGYLRNKKK 58
Db   7  GRGELISAGKCGFRPSDQ--RRCSAQCGGRLPGSRPVSRCVR-----CIC--CSLRASPL 57

QY  59  VCVPRS 64
Db   58  PCQPPS 63

```

```

RESULT 5
US-09-978-825-6458
; Sequence 6458, Application US/09978825
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: US/09/978,825
; CURRENT FILING DATE: 2003-01-29
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 6458
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Propioni acnes
US-09-978-825-6458

```

```

Query Match      15.6%; Score 61; DB 5; Length 66;
Best Local Similarity 31.8%; Pred. No. 17;
Matches 21; Conservative 6; Mismatches 27; Indels 12; Gaps 4;

QY  2  GFGGLGGRGKC---PSNEIFSRCDGRQRCPCPNVVPKPLCIKICAPGCVRLGYLRNKKK 58
Db   7  GRGELISAGKCGFRPSDQ--RRCSAQCGGRLPGSRPVSRCVR-----CIC--CSLRASPL 57

QY  59  VCVPRS 64
Db   58  PCQPPS 63

```

```

RESULT 6
US-10-057-498-6458
; Sequence 6458, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir

```

```

; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 6458
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Propioni acnes
US-10-057-498-6458

```

```

Query Match      15.6%; Score 61; DB 6; Length 66;
Best Local Similarity 31.8%; Pred. No. 17;
Matches 21; Conservative 6; Mismatches 27; Indels 12; Gaps 4;

QY  2  GFGGLGGRGKC---PSNEIFSRCDGRQRCPCPNVVPKPLCIKICAPGCVRLGYLRNKKK 58
Db   7  GRGELISAGKCGFRPSDQ--RRCSAQCGGRLPGSRPVSRCVR-----CIC--CSLRASPL 57

QY  59  VCVPRS 64
Db   58  PCQPPS 63

```

```

RESULT 7
US-10-170-385-239
; Sequence 239, Application US/10170385
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 239
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-170-385-239

```

```

Query Match      15.0%; Score 58.5; DB 6; Length 61;
Best Local Similarity 32.7%; pred. No. 29;
Matches 16; Conservative 6; Mismatches 20; Indels 7; Gaps 4;

QY  1  GFGGLGGRGKCPSNEIFSRCDGRQRCPCPNVVPKPLCIKICAPGCVCR 49
Db   10  GGSACAGSCKCKK---CKCTS-CKKSCCSCCPUG-CAK-CAQGCICK 51

```

```

RESULT 8
PCT-US02-32727-9993
; Sequence 9993, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois

```

```
; SEQ ID NO 445
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-440-068-445
```

```
Query Match      14.7%; Score 57.5; DB 7; Length 61;
Best Local Similarity 36.7%; Pred. No. 37;
Matches 18; Conservative 2; Mismatches 22; Indels 7; Gaps 3;

QY 1 GGFGGLGGRGKCPNFISSRCDGRFCPNVVPKPLCIKICAPGCVCR 49
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10 GGSCTCAGSCKCKECKCTS-CKKSCCSCP-----VGCAR-CAQGCVC 51
```

RESULT 13

```
US-09-724-676-81522
; Sequence 81522, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 81522
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-81522
```

```
Query Match      14.6%; Score 57; DB 5; Length 67;
Best Local Similarity 43.2%; Pred. No. 45;
Matches 16; Conservative 1; Mismatches 16; Indels 4; Gaps 2;

QY 1 GGFGGLGGR-GKCPNFISSRCDGRFCPNVVPK 36
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 16 GGGGGLGGGWCKLE---FPLALGHSNKLCPRVSPFP 49
```

RESULT 14

```
US-09-724-676A-81522
; Sequence 81522, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 81522
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-81522
```

```
Query Match      14.6%; Score 57; DB 5; Length 67;
Best Local Similarity 43.2%; Pred. No. 45;
Matches 16; Conservative 1; Mismatches 16; Indels 4; Gaps 2;

QY 1 GGFGGLGGR-GKCPNFISSRCDGRFCPNVVPK 36
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 16 GGGGGLGGGWCKLE---FPLALGHSNKLCPRVSPFP 49
```

RESULT 15

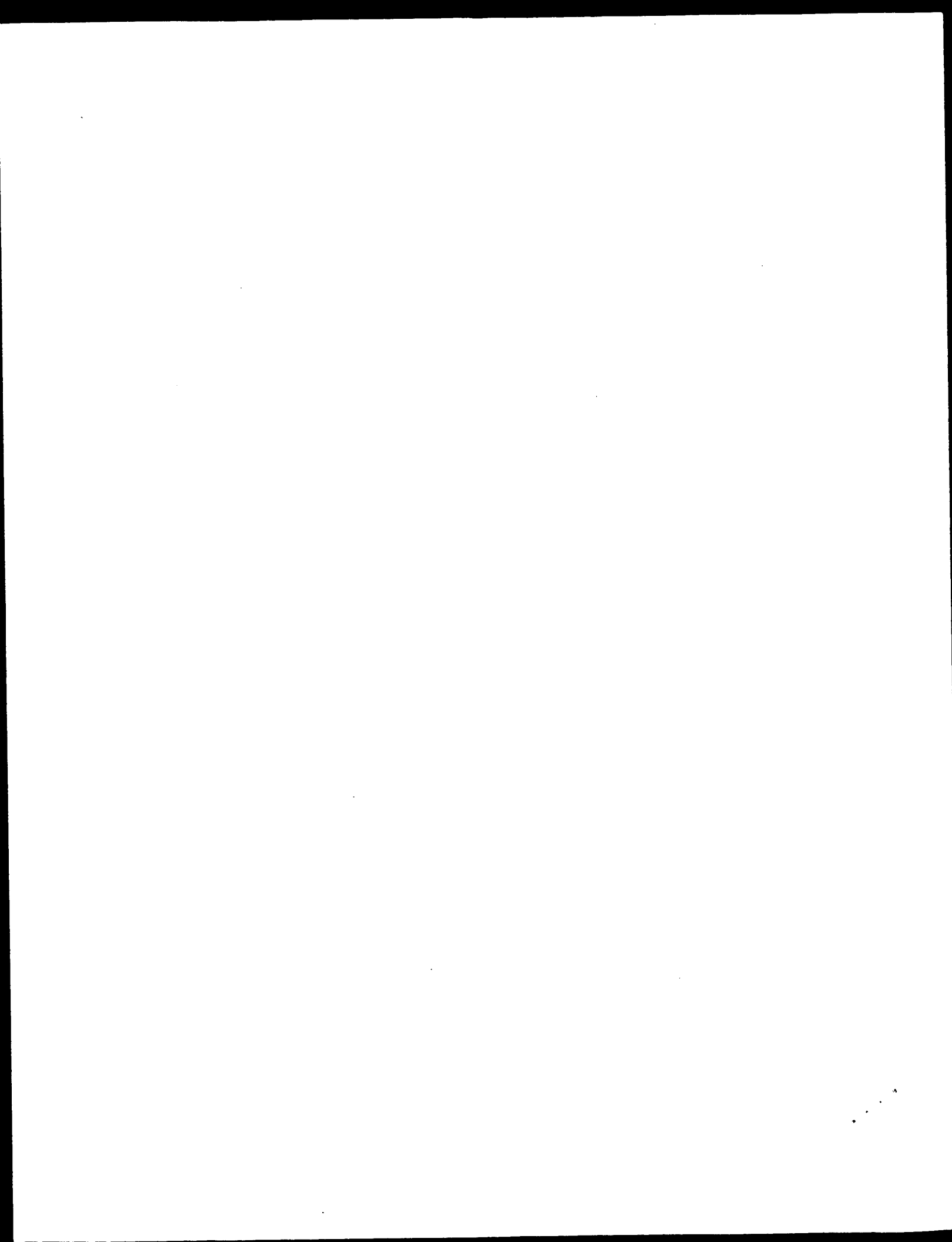
```
US-10-170-385-265
; Sequence 265, Application US/10170385
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
```

```
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 265
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-170-385-265
```

```
Query Match      14.5%; Score 56.5; DB 6; Length 61;
Best Local Similarity 35.0%; Pred. No. 48;
Matches 14; Conservative 6; Mismatches 15; Indels 5; Gaps 4;

QY 10 GKCPNFISSRCDGRFCPNVVPKPLCIKICAPGCVCR 49
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 17 GSCKCKE-CKCTS-CKKSCCSCPVG-CAK-CAQGCICK 51
```

```
Search completed: February 26, 2003, 15:19:50
Job time : 24 secs
```



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: February 26, 2003, 15:10:07 ; Search time 14 Seconds
(without alignments)
140.810 Million cell updates/sec

Title: US-09-506-978-1
Perfect score: 391
Sequence: 1 GGRGGLGGRGKCPSEIFSR.....CRLGYLRNKKKVCVPSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 188167

Minimum DB seq length: 0
Maximum DB seq length: 67

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	82	21.0	61	2	US-08-465-380-61
2	82	21.0	61	2	US-08-486-397-61
3	82	21.0	61	2	US-08-486-399-61
4	82	21.0	61	2	US-08-461-965-61
5	82	21.0	61	2	US-08-634-641-61
6	82	21.0	61	3	US-09-249-471-61
7	82	21.0	61	3	US-09-249-472-61
8	82	21.0	61	3	US-09-249-451-61
9	82	21.0	61	3	US-08-809-455-61
10	82	21.0	61	3	US-09-249-461-61
11	82	21.0	61	3	US-09-249-448-61
12	68	17.4	47	2	US-08-637-759B-400
13	68	17.4	47	3	US-08-871-355A-400
14	68	17.4	47	4	US-09-201-945-400
15	68	17.4	58	2	US-08-465-380-60
16	68	17.4	58	2	US-08-486-397-60
17	68	17.4	58	2	US-08-486-399-60
18	68	17.4	58	2	US-08-461-965-60
19	68	17.4	58	2	US-08-634-641-60
20	68	17.4	58	3	US-09-249-471-60
21	68	17.4	58	3	US-09-249-472-60
22	68	17.4	58	3	US-09-249-451-60
23	68	17.4	58	3	US-08-809-455-60
24	68	17.4	58	3	US-09-249-461-60
25	68	17.4	58	3	US-09-249-448-60
26	62.5	16.0	62	4	US-07-780-717C-5
27	59.5	15.2	61	2	US-08-785-530-3

28 59.5 15.2 61 2 US-09-123-850-3
29 59 15.1 61 2 US-08-785-530-1
30 59 15.1 61 2 US-09-123-850-1
31 57.5 14.7 61 2 US-08-785-530-4
32 57.5 14.7 61 2 US-09-123-850-4
33 57 14.6 10 4 US-09-394-630-20
34 55.5 14.2 38 2 US-08-902-516-47
35 55.5 14.2 45 3 US-08-899-437-11
36 55.5 14.2 45 4 US-09-126-121-11
37 55.5 14.2 46 4 US-08-915-096A-12
38 54.5 13.9 61 2 US-08-785-530-6
39 54.5 13.9 61 2 US-09-123-850-6
40 54 13.8 61 2 US-08-785-530-5
41 54 13.8 61 2 US-09-123-850-5
42 54 13.8 63 2 US-08-369-829A-18
43 52.5 13.4 44 6 5177197-48
44 52 13.3 42 2 US-08-761-248B-15
45 51.5 13.2 48 2 US-08-465-794-3

ALIGNMENTS

RESULT 1
US-08-465-380-61
; Sequence 61, Application US/08465380
; Patent No. 5863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; STATE: Los Angeles
; COUNTRY: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb.
; MEDIUM TYPE: Storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,380
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 61 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Necator americanus

Sequence 3, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 20, Appli
Sequence 47, Appli
Sequence 11, Appli
Sequence 11, Appli
Sequence 12, Appli
Sequence 6, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 18, Appli
Patent No. 5177197
Sequence 15, Appli
Sequence 3, Appli

us-09-506-978-1.closed.raï

Wed Feb 26 15:20:40 2003

US-08-465-380-61

Query Match 21.0%; Score 82; DB 2; Length 61;
Best Local Similarity 30.9%; Pred. No. 0.051;
Matches 17; Conservative 9; Mismatches 19; Indels 10; Gaps 2;

Qy	12	CPSNEIFSR	CDGRCQ	RFQFCP	VVVPKPL	CIKCAP	GCV----	CRGLYLRN	KKKVCV	61
	11:11:	:	:	:	:	:	11	11	11	11:
	11:11:	:	:	:	:	:	11	11	11	11:
Db	4	CPANEWR <th>ECGTPC</th> <th>EPKCNOMP</th> <th>-----</th> <th>DICTMNC</th> <th>IVDV</th> <th>CCCKEGYK</th> <th>RHETGCL</th> <th>53</th>	ECGTPC	EPKCNOMP	-----	DICTMNC	IVDV	CCCKEGYK	RHETGCL	53

RESULT 2

RESOL 2
US-08-486-397-61
: Sequence 61. Application US/08486397

Sequence of: 05/040032
Patent No. 5866542
GENERAL INFORMATION:
APPLICANT: George P. Vlasko, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jaspers,
APPLICANT: Yannick G.J. Gansemaes, Matthew Moyle,

APPLICANT: Peter W. Bergum
TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT

;	TITLE OF INVENTION:	NETA10DD
;	TITLE OF INVENTION:	PROTEIN
;	NUMBER OF SEQUENCES:	357

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.

```

1 21P: 90071
2
3 COMPUTER READABLE FORM:
4 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
5 MEDIUM TYPE: storage
6
7 COMPUTER: IBM Compatible
8 OPERATING SYSTEM: IBM P.C. DOS 5.0
9 SOFTWARE: Word Perfect 5.1
10 CURRENT APPLICATION DATA:
11 APPLICATION NUMBER: US/08/486.397
12

```

FILING DATE: June 3, 1993
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/326,110
 FILING DATE: October 18, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: BIGGS, SUZANNE L.
 REGISTRATION NUMBER: 30,158
 REFERENCE/DOCKET NUMBER: 213/269
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440

TELEX: 07-3310
 INFORMATION FOR SEQ ID NO: 61:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 61 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: *Necator americanus*

Query Match 21.0%; Score 82; DB 2; Length 61;
Best Local Similarity 30.9%; Pred. No. 0.051;
Matches 17: Conservative 9; Mismatches 19; Indels 10; Gaps 2;

QY 12 CPSNEFSRCDGRQRFPCPNVVPKPLCIKCAPGCV-----CRGLYLNRNKKVCV 61
||:||| : | : | : || : | : | : | : | : | : | :
Dk 4 CPANENWPECCCTDCFKCNOBMP-----DICTMNCITDVAVOCKECYKRHETKGCL 53

RESULT 3

US-08-486-399-61
; Sequence 61, Application US/08486399
; Patent No. 5866543

GENERAL INFORMATION:

APPLICANT: George P. Vlasuk, Patric H. Stanssens,
APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
APPLICANT: Yves R. Laroche, Laurent S. Jaspers,
APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
APPLICANT: Peter W. Bergum

1. TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT

```
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 356
```

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.

```

? ZIP: 30011
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5" Diskette, 1.44 MB
? MEDIUM TYPE: storage
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: IBM P.C. DOS 5.0
? SOFTWARE: Word Perfect 5.1
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/486.399
? FILING DATE: June 5, 1995

```

CLASSIFICATION: 330
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/326,110
FILING DATE: October 18, 1994
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/270
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 61:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 61 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Necator americanus
 PS-08-486-399-61

Query Match	21.0%	Score 82:	DB 2:	Length 61:
Best Local Similarity	30.9%	Pred. NO.	0.051:	
Matches	17:	Conservative	9:	Mismatches 19:
		Indels	10:	Gaps 2:

QY 12 CPSNEIFSPCDGRCCRFPCPNVVPKPLCIKICAPGCV-----CRLGYLNRNKKVCV 61
||:||:| | |: | || | : | : ||:||:| | :
Db 4 CPANFEFWRCFGTPECFKNOPMP-----DICTMNCIVDVCCCKEYKRHETGCL 53

RESULT A

RESULT 4
US-08-461-965-61
: Sequence 61 Application US/08461965

Sequence 61; Application 05/06401903

GENERAL INFORMATION:

Patent No. 5872098

APPLICANT: George P. Vlasuk, Patric H. Stanssens,

APPLICANT: Joris H. L. Mensens, Marc J. Lauwereys,

APPLICANT: Yves R. Laroche, Laurent S. Jespers,

APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,

APPLICANT: Peter W. Bergum

TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT

TITLE OF INVENTION: PROTEIN

NUMBER OF SEQUENCES: 356

; Sequence 400, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 400:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-637-759B-400

Query Match 17.4%; Score 68; DB 2; Length 47;
Best Local Similarity 55.6%; Pred. No. 1.1;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 12 CPSNEIFSRCDGRCQRCFC 29
||: : ||||| |
Db 30 CPAGKPLSRCDGRCDEIC 47

RESULT 13
US-08-871-355A-400
; Sequence 400, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 400:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-871-355A-400

Query Match 17.4%; Score 68; DB 3; Length 47;
Best Local Similarity 55.6%; Pred. No. 1.1;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 12 CPSNEIFSRCDGRCQRCFC 29
||: : ||||| |
Db 30 CPAGKPLSRCDGRCDEIC 47

RESULT 14
US-09-201-945-400
; Sequence 400, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,945
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/637,759
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 400:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 47 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-201-945-400

Query Match

Best Local Similarity 17.4%; Score 68; DB 4; Length 47;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 12 CPSNEIFSRCDGRCQRCFC 29
Db 30 CPAGPLSRCDGRCDEIC 47

RESULT 15

US-08-465-380-60
; Sequence 60, Application US/08465380
; Patent No. 5863894
; GENERAL INFORMATION:
; APPLICANT: George P. Vlasuk, Patric H. Stanssens,
; APPLICANT: Joris H.L. Mensens, Marc J. Lauwereys,
; APPLICANT: Yves R. Laroche, Laurent S. Jespers,
; APPLICANT: Yannick G.J. Gansemans, Matthew Moyle,
; APPLICANT: Peter W. Bergum
; TITLE OF INVENTION: NEMATODE-EXTRACTED ANTICOAGULANT
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 356
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,380
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/326,110
; FILING DATE: October 18, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 213/268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 58 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Heligmosomoides polygyrus
US-08-465-380-60

Query Match

Best Local Similarity 17.4%; Score 68; DB 2; Length 58;
Matches 14; Conservative 7; Mismatches 34; Indels 0; Gaps 0;

QY 12 CPSNEIFSRCDGRCQRCFCNVVPKPLCICIKICAPGCVCRGLYLNNKKKVCVPRSKC 66
Db 3 CGPNEEYTECGTCEPKCNEPMDICTLNCIVNVCQCKGFKRGKPGCVAPGPGC 57

Search completed: February 26, 2003, 15:13:22
Job time : 15 secs